
Lexikos at Eighteen: An Analysis

Gilles-Maurice de Schryver, *Department of African Languages and Cultures, Ghent University, Ghent, Belgium; Xhosa Department, University of the Western Cape, Bellville, Republic of South Africa; and TshwaneDJe HLT, Pretoria, Republic of South Africa (gillesmaurice.deschryver@UGent.be)*

Abstract: At eighteen, *Lexikos* became a major player in the field of linguistics, by being awarded an Impact Factor. This article presents a double analysis of the foundation that led to this success. On the one hand a thorough statistical study is undertaken with regard to all contributors and their contributions to *Lexikos*. To this end a metadata database was designed, with the aim to answer the question: 'Who publishes what type of material from where and when?' On the other hand a content analysis is carried out which focuses on the actual topics (i.e. 'keywords') in *Lexikos*. To this end an all-inclusive text corpus containing all the *Lexikos* material was built, with the aim to answer the question: 'What are the major trends in *Lexikos*?'

Keywords: LEXIKOS, LEXICOGRAPHY, METALEXICOGRAPHY, DICTIONARIES, LEXICOGRAPHERS, METADATA DATABASE, TEXT CORPUS, CONTRIBUTIONS, CONTRIBUTORS, AFFILIATIONS, STATISTICS, TRENDS, ENGLISH, AFRIKAANS, BANTU

Samenvatting: *Lexikos* op achttienjarige leeftijd: Een analyse. Op achttienjarige leeftijd werd *Lexikos* een speler van wereldformaat binnen de linguïstiek, door het behalen van een Impact Factor. Dit artikel stelt een dubbele analyse voor van de grondslag die tot dit succes leidde. Aan de ene kant wordt een gedetailleerde statistische studie ondernomen met betrekking tot alle auteurs en hun bijdragen tot *Lexikos*. Daarvoor werd een metadata-databasis ontworpen, waarmee de volgende vraag wordt beantwoord: 'Wie publiceert welk soort materiaal, van waaruit en wanneer?' Aan de andere kant wordt een inhoudsanalyse uitgevoerd die zich toespitst op de feitelijke onderwerpen (m.n. 'sleutelwoorden') in *Lexikos*. Daarvoor werd een alles-inclusief tekstcorpus met al het *Lexikos* materiaal gebouwd, waarmee de volgende vraag wordt beantwoord: 'Wat zijn de voornaamste trends in *Lexikos*?'

Sleutelwoorden: LEXIKOS, LEXICOGRAFIE, METALEXICOGRAFIE, WOORDENBOEKEN, LEXICOGRAFEN, METADATA-DATABASIS, TEKSTCORPUS, BIJDRAGEN, AUTEURS, AFFILIATIES, STATISTIEKEN, TRENDS, ENGELS, AFRIKAANS, BANTOE

1. *Lexikos*: From newborn, to toddler, to adult

In the foreword to the very first issue of the present journal, the genesis of *Lexikos* is explained thus: "a formal discussion on lexicography between the editorial staff of the WAT and a group of prominent linguists got off the ground in Stellenbosch in 1989, [and] the staff is attempting to extend this discussion to include the rest of Africa" (Harteveld 1991: xii). From humble but

enthusiastic beginnings at the Bureau of the *Woordeboek van die Afrikaanse Taal* (WAT) two decades ago, *Lexikos* has been on a straight trajectory, indeed extending its discussion throughout, not only in geographical terms but also in terms of the ever-growing variety and quality of the lexicographic topics discussed. *Lexikos* has now been granted its place amongst the major journals in linguistics, through its inclusion in the Thomson Reuters Web of Science Citation Index since 2005, and the calculation of its impact, by means of an annual Impact Factor, as from last year's issue (De Schryver 2009). It is fitting, therefore, to look back at this point in time — '*Lexikos* at Eighteen' — to see how the toddler turned into a teenager, before it will now continue life as an adult.

Lexikos received some stimuli along the way; the most effective of these was the result of the establishment of the African Association for Lexicography (AFRILEX) in 1995, and the subsequent agreement whereby membership of AFRILEX includes a copy of the annual journal *Lexikos*. As such, *Lexikos* became the official mouthpiece of AFRILEX.

As a newborn, *Lexikos* had to find its way, and learn to walk. The first issue, virtually entirely in Afrikaans, was actually a cross between a journal and a Festschrift (cf. also Du Plessis 2008: xi), and was even reviewed (as if it were a book) for the *International Journal of Lexicography* by the father of modern lexicography: Ladislav Zgusta (1993). From a single file of articles in 1991, the journal morphed into one with multiple subsections, in an attempt to keep abreast of the ever-changing and expanding discipline of lexicography. From Volume 4 to Volume 14, for example, articles were mostly grouped into 'Research Articles' vs. 'Contemplative Articles', in addition to 'Review Articles' and 'Reviews' (of books). In the last volume, Volume 18, by contrast, there are no fewer than twelve categories, from 'Articles' to 'Lexicosoftware' to 'Projects' to 'Lexicosurvey', etc. Other changes include the introduction, in Volume 4, of a 'Review Editor', an 'Advisory Board', an 'Editorial Committee', and a statement of the 'Editorial Objectives' in three languages (Afrikaans, English, and German). Starting with Volume 6, each issue has included 'A Few Words from AFRILEX' (in both Afrikaans and English). 'Instructions to Authors' (again in three languages) first appeared in Volume 8. E-mail addresses for all authors were added for the first time in Volume 11. As of Volume 12, finally, both the 'Editorial Objectives' and the 'Instructions to Authors' have also been presented in a fourth language (French).

The overwhelming majority of the contributions to *Lexikos* are original studies without any prior traces of publication.¹ In order to study the first eighteen years of *Lexikos*, two databases were built. On the one hand a *Lexikos* Metadata Database was designed, in which all the main bibliometric details of each and every contribution were brought together. On the other hand a *Lexikos* Corpus was built, containing the full text of each and every contribution that appeared during the first eighteen years. In what follows, both databases are subjected to various queries, with the results of those queries summarized in tables and figures.

2. The *Lexikos* Metadata Database

2.1 Number of contributions and number of pages

There are a total of 6 587 pages in *Lexikos* 1 (1991) through *Lexikos* 18 (2008), on average 366 pages per year. The actual number of pages per volume may be seen from the line in Figure 1: from about 300 pages in the first few volumes, to about 500 pages in the last few volumes. The number of pages thus continuously rose over the years. In parallel, also the number of contributions rose, as may be seen from the bars in Figure 1: from under 20 contributions in the first two volumes, to about 35 in the last few volumes. There are exactly 543 contributions in all, on average 30 per volume. The Pearson correlation coefficient r between the number of pages and the number of contributions is 0.74, so the fit is not so good. This, on the whole, points to a widely varying average length of the contributions.

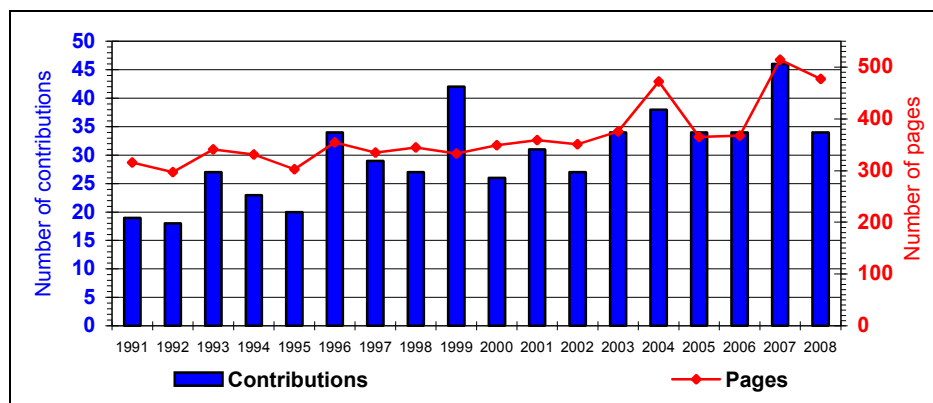
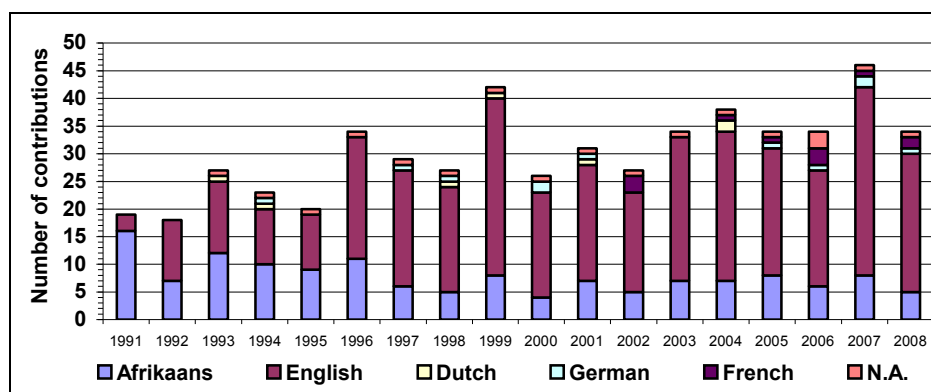
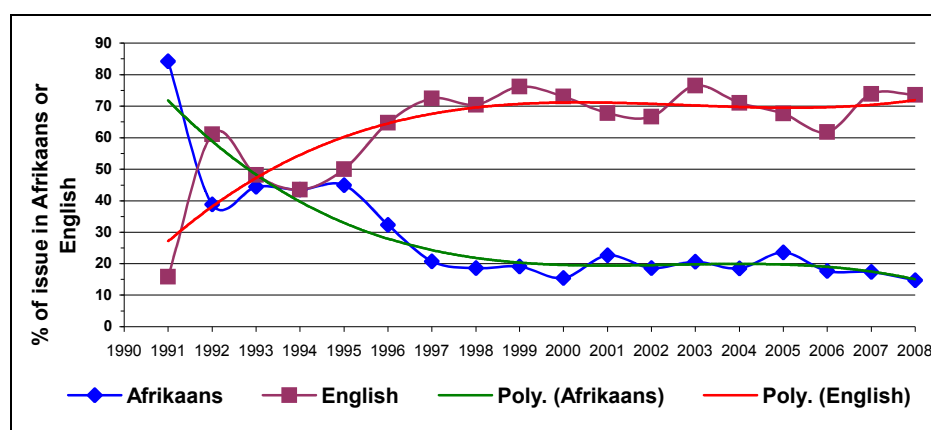


Figure 1: Number of contributions and number of pages in each *Lexikos* issue

Lexikos is a multilingual journal, and contributions in Afrikaans, English, Dutch, German, and French have all been accepted; distributed as seen in Figure 2. Two conclusions may be drawn from Figure 2: (a) Contributions in languages other than Afrikaans or English are by far in the minority; (b) English has rapidly taken over from Afrikaans as the main language of discourse. The latter may also be confirmed from Figure 3, where the number of Afrikaans and English contributions has been expressed as a percentage of the total per *Lexikos* issue. While 84% of the contributions were in Afrikaans in the first issue, this proportion decreases over the years to stabilize to around 20% (cf. the Afrikaans polynomial trend line). In parallel, while only 16% of the contributions were in English in the first volume, this proportion increases over the years to stabilize to about 70% (cf. the English polynomial trend line). (The remainder, $100\% - 20\% - 70\% = 10\%$, accounts for the 'other languages'.)

Figure 2: Distribution of the languages in each *Lexikos* issueFigure 3: Afrikaans vs. English in each *Lexikos* issue (expressed as % of total)

A detailed view of the exact number of contributions and exact number of pages per *Lexikos* issue is shown in Table 1. From that data one sees that, overall, about 26% of the contributions in *Lexikos* are in Afrikaans, 65% in English, 1% in Dutch, 2% in German, and another 2% in French. A further 3% cannot be linked to particular languages, as they consist of lists of publications. Once English and Afrikaans had been 'established' as the discourse languages of *Lexikos*, the next (logical) language that was added was Dutch (from 1993 until 2004), followed by German (since 1994), and then French (since 2002).

Looking at the three main 'peaks' in Figures 1 and 2, finally, one can observe that all three are the results of one-time 'additions': in 2007 a special section was added, called 'Lexicofocus', in which the lexicographic activities at ALRI (in Harare, Zimbabwe) were surveyed; in 2004 a special section on 'Learners' Dictionaries' was added; while in 1999 the full proceedings of a meeting, 'Lexicography in a Multilingual South Africa', was added. Page-wise,

especially the 2007 and 2004 additions had a great effect (by contrast, the 1999 papers were all short); expressed in number of contributions, all three additions of course had a considerable effect.

Table 1: Detail of the number of contributions and number of pages in each *Lexikos* issue (with A = Afrikaans, E = English, NL = Dutch, DE = German, FR = French)

Year	A	%	E	%	NL	%	DE	%	FR	%	N.A.	%	Items	Roman pages	Arabic pages	Sum pages
1991	16	84.21	3	15.79	–	–	–	–	–	–	–	–	19	xix	297	316
1992	7	38.89	11	61.11	–	–	–	–	–	–	–	–	18	x	287	297
1993	12	44.44	13	48.15	1	3.70	–	–	–	–	1	3.70	27	x	331	341
1994	10	43.48	10	43.48	1	4.35	1	4.35	–	–	1	4.35	23	xi	320	331
1995	9	45.00	10	50.00	–	–	–	–	–	–	1	5.00	20	xiii	290	303
1996	11	32.35	22	64.71	–	–	–	–	–	–	1	2.94	34	xvi	339	355
1997	6	20.69	21	72.41	–	–	1	3.45	–	–	1	3.45	29	xiv	321	335
1998	5	18.52	19	70.37	1	3.70	1	3.70	–	–	1	3.70	27	xiv	331	345
1999	8	19.05	32	76.19	1	2.38	–	–	–	–	1	2.38	42	xv	318	333
2000	4	15.38	19	73.08	–	–	2	7.69	–	–	1	3.85	26	xiv	335	349
2001	7	22.58	21	67.74	1	3.23	1	3.23	–	–	1	3.23	31	xv	344	359
2002	5	18.52	18	66.67	–	–	–	–	3	11.11	1	3.70	27	xvi	335	351
2003	7	20.59	26	76.47	–	–	–	–	–	–	1	2.94	34	xv	361	376
2004	7	18.42	27	71.05	2	5.26	–	–	1	2.63	1	2.63	38	xv	457	472
2005	8	23.53	23	67.65	–	–	1	2.94	1	2.94	1	2.94	34	xv	350	365
2006	6	17.65	21	61.76	–	–	1	2.94	3	8.82	3	8.82	34	xiii	355	368
2007	8	17.39	34	73.91	–	–	2	4.35	1	2.17	1	2.17	46	xvi	498	514
2008	5	14.71	25	73.53	–	–	1	2.94	2	5.88	1	2.94	34	xiii	464	477
	141	25.97	355	65.38	7	1.29	11	2.03	11	2.03	18	3.31	543	ccliv	6 333	6 587

2.2 Types of contributions

The second series of queries one can submit to the *Lexikos* Metadata Database are those concerning the types of contributions. Using the *Lexikos*-internal categorization, the data shown in Table 2 is obtained.

Table 2: Types of contributions in *Lexikos* (according to *Lexikos*-internal categories)

Type	N	%	Type	N	%
Review	111	20.44	Meeting	14	2.58
Article	106	19.52	Lexicofocus	10	1.84
Contemplative Article	68	12.52	Lexicosoftware	6	1.10
Editorial Material	64	11.79	Lexiconews	2	0.37
Research Article	49	9.02	Lexicotribute	2	0.37
Lexiconote(s)	26	4.79	Report	2	0.37
Review Article	23	4.24	Lexicobibliography	1	0.18
Project(s)	20	3.68	Lexicohonour	1	0.18
Lexicovaria	19	3.50	Lexicosurvey	1	0.18
Publication Announcements	18	3.31		543	100.00

Obviously, some of the categories seen in Table 2 should be collapsed, as they are the result of diverging and converging attempts to categorize the various contributions over the years. In regrouping the types, we were inspired by the Web of Science categorization on the one hand, as well as the section on 'type and content of articles' as found in the *Lexikos* 'Instructions to Authors' (where the main dichotomy is one between articles subjected to strict anonymous refereeing vs. those that are not). Doing so, we obtained seven categories: ARTICLE (i.e. Article, Contemplative Article, Research Article, Lexicofocus, Lexicosoftware, and Lexicosurvey), REVIEW ARTICLE (idem), BOOK REVIEW (i.e. Review), PROCEEDINGS PAPER (i.e. Meeting), OTHER (i.e. Lexiconote(s), Project(s), Lexicovaria, Lexiconews, Lexicotribute, Report, Lexicobibliography, and Lexicohonour), EDITORIAL MATERIAL (idem), and PUBLICATION ANNOUNCEMENTS (idem). The ARTICLE and REVIEW ARTICLE sections are subjected to strict anonymous refereeing; the others must all meet the requirements of academic writing. The resulting distribution is tabulated in Table 3.

Table 3: Types of contributions in *Lexikos* (inspired by Web of Science categories)

Type	N	%
Article	240	44.20
Review Article	23	4.24
Book Review	111	20.44
Proceedings Paper	14	2.58
Other	73	13.44
Editorial Material	64	11.79
Publication Announcements	18	3.31
	543	100.00

As may be seen, as many as 44% of all contributions to *Lexikos* are true articles, a further 4% are review articles, and 20% are book reviews.²

Given that an analogous metadata study was performed for the *International Journal of Lexicography* (IJL) for the period January 1988 up to December 2008 (cf. De Schryver 2009a), a comparison of the values is instructive. In IJL, one finds 31% articles, 2% review articles, and 43% book reviews. While, in quantitative terms, IJL tends to focus on book reviews, the reverse is true for *Lexikos*, where most attention goes to articles and review articles.

2.3 Contributors and their contributions

A total of 555 names (i.e. name tokens) appear under the 543 *Lexikos* contributions. These two values are very close, and indeed, most contributions are authored by just one scholar, as may be seen from Table 4. As many as 83% of all contributions to *Lexikos* are written by a single author, a further 12% by two. There are very few with three, and just two with four — the maximum. By

contrast, in IJL one finds contributions with up to 12 authors. The single- and co-authored publication values in IJL are, however, very similar to those in *Lexikos*: 81% for sole authorship, 13% for co-authorship.

Table 4: Author position in the *Lexikos* Metadata Database

Author position	N	%
1	459	82.70
2	64	11.53
3	10	1.80
4	2	0.36
(no author)	20	3.60
	555	100.00

The 555 names that appear under the 543 contributions may be collapsed into 260 different (unique) names.³ In other words, a total of 260 different scholars have written material for *Lexikos* over the past eighteen years. The distribution of the number of contributions for each of those is as shown in Table 5.

Table 5: Number of contributions per author in the *Lexikos* Metadata Database

# contributions	N	%	# contributions	N	%
1	171	65.77	10	1	0.38
2	42	16.15	12	1	0.38
3	12	4.62	14	1	0.38
4	10	3.85	17	1	0.38
5	6	2.31	18	1	0.38
6	3	1.15	20	1	0.38
7	3	1.15	32	1	0.38
8	2	0.77	32	1	0.38
9	3	1.15		260	100.00

Two thirds (66%) of the contributors to *Lexikos* only contributed once, 16% twice, 5% thrice, etc. In contrast, in IJL as many as three quarters (74%) contributed only once, 11% twice, 6% thrice, etc. *Lexikos*, then, has far more repeat contributors than IJL. However, the maximum number of contributions for a single contributor stands at around thirty for both journals.

Studying who is behind the high number of contributions, one obtains the data shown in Tables 6 and 7. In Table 6 the focus is on the peer-reviewed contributions: articles and review articles (as regrouped above). The top five here is: Gouws, De Schryver, Prinsloo, Wiegand, and Louw. In Table 7 the focus is on the book reviews. The top five there is: Abecassis, Smit, Louw, Nkabinde, and Swanepoel. The overall top five brings together the *Lexikos* Editor and the four AFRILEX Presidents so far, rather surprisingly in their tenure order: Gouws (with 32 contributions), Du Plessis (32), Prinsloo (20), Alberts (18), and De Schryver (17). Of course, the high overall ranking for the first four is helped by the editorial material each wrote.⁴

Table 6: Top contributors to articles and review articles in *Lexikos* (with A = article and review article, B = book review, P = proceedings paper, O = other, E = editorial material; -, ✓ or ☑ = 0, 1 or more than 1 article or review article per year)

#	Author	Lexikos 1 (1991) through 18 (2008)	A	B	P	O	E	N
1	Gouws RH	✓-☑✓-☑✓✓✓✓✓✓✓✓✓✓✓✓	20	2	-	-	10	32
2	De Schryver G-M	-----✓✓☑✓✓☑☑☑☑	16	1	-	-	-	17
3	Prinsloo DJ	- ✓-----✓✓- ☑✓☑✓✓- ✓--	11	-	1	-	8	20
4	Wiegand HE	-----✓--- ✓- ✓✓✓✓✓✓✓	9	-	-	-	-	9
5	Louw PA	-----✓✓☑☑-----✓✓--	8	4	-	-	-	12
6	Otto AN	-- ✓✓-----✓✓- ✓- ✓-- ✓--	7	1	-	-	-	8
7	Swanepoel PH	✓✓--- ✓----- ✓----- ✓-- ✓	6	3	-	-	-	9
8	Carstens A	---- ☑☑- ✓- ✓-----	6	1	-	-	-	7
9	Alberts M	- ✓----- ✓✓- ✓✓✓-----	6	-	-	4	8	18
10	Botha W	✓✓✓- ✓✓----- ✓-----	6	-	-	1	-	8
11	Taljar E	----- ✓✓- ☑-- ☑--	6	-	-	-	-	6
12	Mavoungou PA	----- ✓✓✓- ✓✓-	5	1	-	1	-	7
13	Chabata E	----- ✓- ✓- ✓✓----- ✓-	5	-	-	1	-	6
14	Tarp S	----- ✓--- ☑-- ✓✓	5	-	-	-	-	5
15	Smit M	----- ✓- ✓✓✓-----	4	6	-	-	-	10
16	Beyer HL	----- ✓- ✓----- ✓✓-	4	1	-	-	-	5
	Feinauer I	----- ✓----- ☑✓	4	1	-	-	-	5
18	Lombard FJ	✓✓✓✓-----	4	-	-	-	-	4
	Mdee JS	--- ✓--- ✓- ✓----- ✓-	4	-	-	-	-	4
20	Potgieter L	----- ✓✓✓- ✓✓✓	3	1	-	-	-	4
21	Al-Ajmi H	----- ✓----- ✓✓	3	-	-	1	-	4
	Bergenholtz H	----- ✓--- ✓✓	3	-	-	1	-	4
	Hadebe S	----- ✓- ✓--- ✓-	3	-	-	1	-	4
	Hartmann RRK	-- ✓-- ✓----- ✓--	3	-	-	1	-	4
25	Gauton R	----- ✓----- ✓✓	3	-	-	-	-	3
	Joffe D	----- ☑✓--	3	-	-	-	-	3
	Kotzé EF	- ✓----- ✓- ✓-----	3	-	-	-	-	3

Table 7: Top contributors to book reviews in *Lexikos*

#	Author	Lexikos 1 (1991) through 18 (2008)	A	B	P	O	E	N
1	Abecassis M	----- ✓☑☑- ✓	2	7	-	-	-	9
2	Smit M	----- ✓✓- ✓☑✓	4	6	-	-	-	10
3	Louw PA	----- ✓☑----- ✓--	8	4	-	-	-	12
4	Nkabinde AC	-- ☑-----	1	4	1	-	-	6
5	Swanepoel PH	----- ✓----- ☑-----	6	3	-	-	-	9
6	Heliel MH	- ✓----- ✓- ✓-----	2	3	-	-	-	5
	Prinsloo AF	----- ✓----- ✓--- ✓-	2	3	-	-	-	5
8	Du Plessis JCMD	----- ✓--- ✓- ✓--	1	3	2	2	24	32
9	Ridge E	-- ✓✓----- ✓-----	-	3	-	-	-	3

Also shown in Tables 6 and 7 are the particular volumes of *Lexikos* in which the articles and review articles on the one hand, and the book reviews on the other, appeared. From these distributions across time, one may conclude that the

repeat contributors are also 'faithful' contributors to *Lexikos*, in the sense that once they start contributing, they also tend to continue to contribute over the years. In Table 6 there are only four uneven distributions (Carstens, Smit, Lombard, and Kotzé), and in Table 7 three (Nkabinde, Heliel, and Ridge). Note, further, that the two tables are complementary for three contributors (Smit, Louw, and Swanepoel), whereby a 'gap' in one category actually means that efforts have been put in another category (this is especially the case for Smit, who switched her attention from articles to book reviews over the years). Given that the ticks also tend to cluster near the right-hand side of the time scale, one may further conclude that *Lexikos* attracts ever more repeat contributors. This finding is clearly a very welcome one, as it bodes well for the future of *Lexikos*.

2.4 Continents, countries, cities and affiliations of contributors

A total of 649 affiliations (i.e. affiliation tokens) accompany the 543 *Lexikos* contributions. These will now be analysed, starting with a world view down, zooming in on the most productive region(s) at each step.

Although contributions are submitted from every corner of the world, there is a clear bias towards researchers who work from Africa, as seen in Table 8. As many as three quarters (74%) come from Africa, and 18% from Europe. Tiny numbers come from the Americas, the Middle-East, Asia and Australasia.

Table 8: Continents of contributors in *Lexikos*

Continent	N	%
Africa	478	73.65
Europe	116	17.87
Americas	12	1.85
Middle-East	10	1.54
Asia	7	1.08
Australasia	5	0.77
(no affiliation)	21	3.24
	649	100.00

Within Africa, over 86% of the contributions come from South Africa, 6% from Zimbabwe, 2% from Namibia, etc. See Table 9. The top three for Europe is: Belgium, Germany, and the UK, each with about one-fifth of the contributions from Europe. For the Americas, 10 come from the USA and 2 from Canada; for the Middle-East, 8 come from Kuwait and 2 from Egypt; for Asia, 4 come from Hong Kong and 3 from Japan; and for Australasia, 5 come from Australia.

With about 200 countries in the world, that means that about 15% is represented in *Lexikos*. With about 50 countries each in Africa and Europe, about 20% of the African as well as 20% of the European countries are represented. The other continents cannot really be said to be represented; rather, they are sampled.

Table 9: Countries of contributors in *Lexikos*

#	AFRICA	N	%	#	EUROPE	N	%	#	AMERICAS	N	%
1	South Africa	413	86.40	1	Belgium	25	21.55	1	USA	10	83.33
2	Zimbabwe	31	6.49	2	Germany	24	20.69	2	Canada	2	16.67
3	Namibia	11	2.30	3	UK	23	19.83			12	100.00
4	Tanzania	6	1.26	4	Denmark	16	13.79	#	MIDDLE-EAST	N	%
5	Gabon	5	1.05		Netherlands	16	13.79	1	Kuwait	8	80.00
6	Botswana	3	0.63	6	Poland	4	3.45	2	Egypt	2	20.00
	Kenya	3	0.63	7	Sweden	3	2.59			10	100.00
8	DR Congo	2	0.42	8	Norway	2	1.72	#	ASIA	N	%
	Ghana	2	0.42	9	Finland	1	0.86	1	Hong Kong	4	57.14
10	Nigeria	1	0.21		France	1	0.86	2	Japan	3	42.86
	Zambia	1	0.21		Switzerland	1	0.86			7	100.00
		478	100.00			116	100.00	#	AUSTRALASIA	N	%
								1	Australia	5	100.00
										5	100.00

Zooming in on South Africa, where 64% of all *Lexikos* contributions come from, one obtains the breakdown in terms of cities shown in Table 10.

Table 10: South African cities of contributors in *Lexikos*

#	City	N	%	#	City	N	%
1	Stellenbosch	190	46.00	13	Potchefstroom	3	0.73
2	Pretoria	122	29.54	14	Alice	2	0.48
3	Johannesburg	18	4.36		Isipingo	2	0.48
4	Bellville	14	3.39		Mankweng	2	0.48
5	Polokwane	11	2.66	17	Jeffreys Bay	1	0.24
6	Bloemfontein	9	2.18		Letaba	1	0.24
7	Port Elizabeth	8	1.94		Linden	1	0.24
8	Grahamstown	6	1.45		Mafikeng	1	0.24
	KwaDlangezwa	6	1.45		Thornville	1	0.24
10	Cape Town	5	1.21		Upington	1	0.24
11	Durban	4	0.97		Wellington	1	0.24
	Soshanguve	4	0.97			413	100.00

Within South Africa, 46% emanate from Stellenbosch; the second spot is for Pretoria, with 30% of all contributions; the third for Johannesburg, with just 4%. Zooming in on Pretoria, as done in Table 11, one sees that several groups at the University of Pretoria as well as at the University of South Africa are very actively contributing to *Lexikos*, with one third (33%) from the Department of African Languages at the University of Pretoria alone.

Table 12, lastly, zooms in on Stellenbosch. The two biggest contributing centres here are the Department of Afrikaans and Dutch at the University of Stellenbosch with 39%, and of course, with 45%, the Bureau of the WAT itself — the birthplace of *Lexikos*, and also where the journal is still being edited and printed.

Table 11: *Lexikos* contributors from Pretoria

Affiliation	N	%
<i>University of Pretoria</i>	54	44.26
Department of African Languages	40	32.79
Department of Computer Science	2	1.64
Unit for Language Skills Development	2	1.64
Department of Afrikaans	1	0.82
Northern Sotho NLU – Pretoria Branch	1	0.82
School for Languages	1	0.82
"University of Pretoria"	7	5.74
<i>University of South Africa</i>	25	20.49
Department of Afrikaans (and Theory of Literature)	9	7.38
Department of African Languages	7	5.74
Department of Linguistics	3	2.46
Department of Art History, Visual Arts and Musicology	2	1.64
Department of Criminal and Procedural Law	1	0.82
Department of Criminology	1	0.82
Department of Public Administration	1	0.82
Department of Social Work	1	0.82
TshwaneDJe HLT	14	11.48
Pan South African Language Board (PanSALB)	11	9.02
Department of Arts and Culture	6	4.92
National Terminology Services, Department of National Education	5	4.10
(<i>independent scholar</i>)	5	4.10
Translate.org.za	1	0.82
Unit for African Languages, Human Sciences Research Council	1	0.82
	122	100.00

Table 12: *Lexikos* contributors from Stellenbosch

Affiliation	N	%
<i>University of Stellenbosch</i>	101	53.16
Department of Afrikaans and Dutch	75	39.47
Department of Music	11	5.79
SU Centre for Language and Speech Technology (SU-CLaST)	2	1.05
Department of African Languages	1	0.53
Department of Curriculum Studies	1	0.53
Department of Didactics	1	0.53
Department of English	1	0.53
Department of Logistics	1	0.53
Law Faculty	1	0.53
Medicine Faculty	1	0.53
Unit for Afrikaans, Language Centre	1	0.53
Unit for isiXhosa, Language Centre	1	0.53
"University of Stellenbosch"	4	2.11
Bureau of the WAT	86	45.26
(<i>independent scholar</i>)	3	1.58
	190	100.00

3. The *Lexikos* Corpus

In order to transcend the mere bibliometric details discussed in Section 2, a corpus was built consisting of every single contribution in *Lexikos* so far, except for the 18 lists of 'Publication Announcements'. It thus contains $543 - 18 = 525$ files, and after conversion to plain text the size of that corpus is 17MB. The *Lexikos* Corpus can be loaded into and queried with a corpus tool, just like any other corpus. We used WordSmith Tools (Scott 2009). The total size of the *Lexikos* Corpus is about 2.5 million running words, with about 100 000 distinct words (2 493 250 tokens; 103 069 types). A subsection of this full corpus was extracted, namely all English contributions, 355 files in all (as there are 170 non-English files, cf. Table 1), good for 1 744 536 tokens and 65 485 types. As a comparison, the English part of the IJL corpus (covering the period January 1988 up to December 2008) contains 592 files, good for 3 042 282 tokens and 111 903 types (cf. De Schryver 2009a).

Why undertake to build such a corpus? The more traditional approach to study journal contents is to assign keywords to each contribution, and then to draw up various tables that show where (i.e. in which issues, and thus which years) certain topics are being discussed. From such manually drawn-up tables one can then attempt to deduce certain trends. Such an approach has been followed for *Lexikos* before, in an M.A. dissertation (De Schryver 1999: 16-17). Although manual approaches are most definitely valuable, in today's day and age the use of automated approaches with which unlimited amounts of data may be manipulated, and with which that data may be processed and reprocessed ad infinitum, is fast becoming the norm. Querying the *Lexikos* Corpus thus enables one to 'look' at the full contents of *Lexikos* all at once.

The main 'tool' that will be used below to do so is WordSmith's KeyWords function, a tool that automatically extracts words that 'stand out' in a first corpus by means of a comparison with the relative frequencies of those same words in a much larger reference corpus. The reference corpus used is the 100-million-word BNC World Edition.⁵ Moreover, in order to uncover trends, the English section of the *Lexikos* Corpus was partitioned in three, as follows:

- Phase 1: 1991–1996 (six years, 69 files)
- Phase 2: 1997–2002 (six years, 130 files)
- Phase 3: 2003–2008 (six years, 156 files)

In Addenda 1 to 3, the top 300 keywords in each of these three *Lexikos* sub-corpora are listed. In analysing those, it is important to recall that the *full text* of each and every (English language) contribution is included: title, affiliation(s), abstract, keywords, entire article text, endnotes, references, and addenda — as applicable. Given that all these levels are included, one obtains a better picture of what is truly happening in *Lexikos*, especially compared to a study that would merely take the stated keywords into account, for example.

The Addenda are a real source of information, and only a few selected aspects can be touched upon within the space of a single journal article. Each of the lists of keywords contains at least seven types of terms:

- Academic speak (cf. next paragraph)
- Case study terms (cf. next paragraph)
- Non-English words (cf. next paragraph)
- Dictionaries (cf. Section 3.1)
- Proper names (cf. Section 3.2)
- Languages (cf. Section 3.3)
- True keywords (cf. Section 3.4)

'Academic speak' refers to the academic formulae, sentence and thought connectors, as well as abbreviations typically used in academic writing. These include: according, also, as such, based, cf., Ed., Eds., et al., etc., given, in, is, namely, non-*x*, of, or, presented, regard, should, specific, therefore, these, thus, various, and vs. 'Case study terms' are words that have been used in particular case studies and are by consequence very frequent in (some sections of) the corpus. Examples include: kaffir/kaffer, oneself, manifestable, considerate, thoughtful, etc. 'Non-English words' are often also the result of particular case studies or of the text in foreign examples used, such as: *bohlale, dipuku, ka, ke, le, morutiši, mosadi, mpša, na, sa, wa, ya, ...* in Northern Sotho, or *achtung, angst, freude, gefühl, überwinden, ...* in German. These are not particularly interesting for the purposes of the current study. The other categories are, and are looked into in the following sections.

3.1 Dictionaries in *Lexikos*

Extracting the dictionary names from the Addenda one obtains Table 13.

Table 13: Dictionaries in the three *Lexikos* sub-corpora

Phase 1 (1991–1996)		Phase 2 (1997–2002)		Phase 3 (2003–2008)	
Keyword	Keyness	Keyword	Keyness	Keyword	Keyness
WAT	3039.17	WAT	3318.27	WOORDEBOEK	2994.28
WOORDEBOEK	1384.41	WOORDEBOEK	2667.21	WAT	2895.51
AFRIKAANSE	1181.81	AFRIKAANSE	1515.44	ISN	1624.42
OXFORD	411.96	COBUILD	850.31	AFRIKAANSE	1530.81
WNT	360.16	DICTIONNAIRE	706.04	DURAMAZWI	1396.02
COBUILD	330.31	DURAMAZWI	575.86	DICTIONNAIRE	1164.80
ECD	300.46	DGC	538.35	OXFORD	1120.67
DICTIONNAIRE	270.04	RECHISHONA	515.24	COBUILD	1083.54
OED	267.87	MACQUARIE	492.82	RECHISHONA	866.49
		GROOT	458.70	LONGMAN	829.72
		ISN	444.37	ISICHAZAMAZWI	625.79
		VERKLARENDE	434.42		

The picture is rather stable, both in terms of number of keywords and keyness values. The WAT (*Woordeboek van die Afrikaanse Taal*) and COBUILD dictionaries appear throughout. During Phase 1, the WNT (*Woordenboek der Nederlandse Taal*), the ECD (*Explanatory Combinatorial Dictionary*), and the OED (*Oxford English Dictionary*) also receive considerable attention. Moving to Phase 2, Oxford dictionaries disappear out of the top keyword list, while three Zimbabwean dictionaries enter — *Duramazwi reChiShona*, DGC (*Duramazwi Guru reChiShona*), and ISN (*Isichazamazwi SesiNdebele*) —, as well as the *Macquarie Dictionary* from Australia, the *Groot Noord-Sotho woordeboek*, the *Groot Woordeboek/Major Dictionary*, and the HAT (*Verklarende Handwoordeboek van die Afrikaanse Taal*). This state is consolidated in Phase 3. Throughout, a variety of French dictionaries is also referred to.

3.2 Lexicographers in *Lexikos*

There are two types of names in the corpus: personal and place names. Extracting all personal names from the Addenda one obtains Table 14.

Table 14: Lexicographers in the three *Lexikos* sub-corpora

Phase 1 (1991–1996)		Phase 2 (1997–2002)		Phase 3 (2003–2008)	
Keyword	Keyness	Keyword	Keyness	Keyword	Keyness
HARTMANN	1301.68	WIEGAND	7095.03	GOUWS	3321.64
WIEGAND	984.95	GOUWS	2121.63	SCHRYVER	3163.63
HAUSMANN	766.13	PRINSLOO	1665.49	WIEGAND	3051.72
BLOUNT	686.81	HAUSMANN	1543.78	BERGENHOLTZ	2464.73
BRANFORD	559.64	WIEGAND'S	1353.79	TARP	2134.20
ZGUSTA	495.23	SCHRYVER	1261.29	PRINSLOO	2133.77
NADEN	478.79	ALBERTS	682.21	HARTMANN	1518.58
GOUWS	472.71	GILLES	653.54	HESELENG	1232.34
OSSELTON	270.12	BERGENHOLTZ	636.47	PHILIPPSON	1213.09
TONO	269.36	ZGUSTA	626.37	GILLES	1105.95
ILSON	258.87	KRIEL	599.05	HAUSMANN	1047.65
MEL'CUK	247.61	HARTMANN	592.92	CHIMHUNDU	991.65
		TARP	585.90	SCHOLTZ	983.35
		CLUVER	575.86	ZGUSTA	847.23
		LOUW	557.01	MAVOUNGOU	808.72
		LOUWRENS	545.55	HADEBE	711.44
		WYK	541.63	BOSMAN	683.80
		SMIT	525.56	SVEN	683.58
		SAGER	474.41	NIENABER	655.69
		KONERDING	444.52	HÜLLEN	654.68
				MAROLE	606.54
				TOIT	599.28
				RAPONDA	548.77
				MAURICE	547.90
				HENNING	545.84
				JOFFE	544.51

One firstly observes that the number of lexicographers competing for keyword status grows with each new phase (those scholars thus displace true lexicographic keywords, cf. Section 3.4). While the 'big shots' Hartmann, Wiegand, Hausmann and Zgusta are represented throughout, the overpowering presence of Wiegand in Phase 2 (note his keyness value there!) is astonishing. If anything, this points to *Lexikos*'s openness towards metalexicography. With each volume of *Lexikos*, Gouws is gathering more and more attention; see how he rises through the phases. Important here is that his keyness values also increase at each step. The fact that Zgusta seems to go down from one phase to the next is thus misleading: his keyness values too, are increasing at each step; cf. the highlights in Table 14.

3.3 Languages in *Lexikos*

Extracting the languages from the Addenda one obtains Table 15.

Table 15: Languages in the three *Lexikos* sub-corpora

Phase 1 (1991–1996)		Phase 2 (1997–2002)		Phase 3 (2003–2008)	
Keyword	Keyness	Keyword	Keyness	Keyword	Keyness
AFRIKAANS	11942.42	AFRICAN	7485.60	ENGLISH	9489.79
ENGLISH	8755.20	ENGLISH	6987.90	SHONA	8319.34
AFRICAN	4861.51	AFRIKAANS	5313.14	NDEBELE	7895.39
DUTCH	3783.88	SOTHO	3946.56	AFRIKAANS	6101.93
SWAHILI	2202.69	HAUSA	3424.94	AFRICAN	5608.01
YIDDISH	1664.82	SEPEDI	3212.81	SOTHO	4607.16
ZULU	1365.02	SHONA	3058.44	SWAHILI	3051.10
JEWISH	1310.92	ZULU	1648.46	SESOTHO	3029.74
ARABIC	1151.52	NDEBELE	1485.13	LEBOA	2512.87
XHOSA	1144.14	XHOSA	1335.53	ZULU	2401.48
NETHERLAND	1057.53	SWAHILI	1289.27	SETSWANA	1778.25
SOTHO	945.81	SESOTHO	1122.08	FRANÇAIS	1771.51
BENG	826.54	FRANÇAIS	990.08	VENDA	1409.96
MALAYU	810.37	ARABIC	879.79	BANTU	1169.46
AMERICAN	692.96	BANTU	834.20	ISIXHOSA	982.02
KISWAHILI	480.59	CILUBÀ	788.02	SHANGANI	953.14
HOTTENTOT	478.79	ISIXHOSA	616.27	TSHIVENDA	943.51
MUSLIM	478.49	ISIZULU	606.17	SEPEDI	905.00
ISIXHOSA	472.71	DUTCH	519.55	YILUMBU	895.37
BANTU	441.47	SETSWANA	518.29	KHWE	779.84
ASHKENAZIC	384.80	NOORD	515.43	DEUTSCH	690.33
SHONA	358.27	NORTHERN	456.66	HAUSA	685.63
KISI	337.65	VENDA	454.28	KISWAHILI	669.92
NORTHEASTERN	331.40	SWEDISH	442.12	YIPUNU	616.17
MALAYO	292.63			NGUNI	605.85
AUSTRALIAN	261.75			DANISH	591.96
OTJIHERERO	247.61			ZEZURU	577.66
NGUNI	239.39				

Note that, with 'Languages in *Lexikos*', we are not referring to the languages in which the contributions are written (this aspect was dealt with above, cf. Figure 2), but to the languages that are the objects of the investigations. Those languages are overwhelmingly English and Afrikaans, and a good number of Bantu languages. Most key Bantu languages are the (official) South African ones, in addition to: Swahili (Kiswahili), Shona (with dialect Zezuru), Cilubà, Otjiherero, Shangani, Yilumbu, and Yipunu. African non-Bantu languages include: Beng, 'Hottentot', Kisi, Hausa, and Khwe. Also key are: Dutch, New Netherland Dutch, and Cape Muslim Afrikaans; (Northeastern) Yiddish, Jewish, and Arabic; American English, Ashkenazic English, and Australian English; French, Swedish, German, and Danish; as well as Malayu/o.

Although this is an interesting distribution, one cannot escape the fact that these language lists are biased towards the official South African languages, so this warrants a more in-depth study. To do so, eighteen different sub-corpora were built, one for each *Lexikos* volume, from which all non-English contributions were again subtracted. Comparing each of these eighteen sub-corpora with the BNC, keyness values can now be calculated 'per year', and set out graphically.⁶ In Figure 4 this is done for the keywords Afrikaans, English and Bantu.

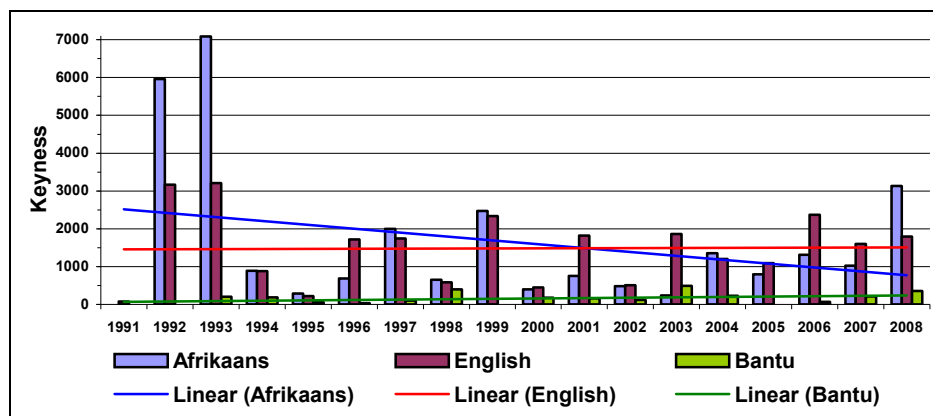


Figure 4: Keyness values for *Afrikaans*, *English* and *Bantu* in each *Lexikos* issue

On the whole, the keyness value of Afrikaans is decreasing over the years, that of English is mostly stable, while that of Bantu rises only slightly. The term 'Bantu' being stigmatized in South Africa, even with reference to the language family, it is not entirely surprising that the term is not really picking up speed.

In order to unravel the real trends, then, one needs to go down to the language level. In Figure 5 this is done for the so-called Sotho group of southern Bantu languages. Note, however, that 'Sotho' can refer to the group itself, but also to two of its members (Northern Sotho and Southern Sotho), while 'Sesotho' can refer to Southern Sotho or even Northern Sotho when that language is referred to in Northern Sotho itself, viz. as 'Sesotho sa Leboa'. Sepedi, although

actually but one of the dialects of Northern Sotho, normally unambiguously refers to Northern Sotho. The third member of the Sotho family is unproblematic: it is either referred to as Tswana (in English) or Setswana (in Tswana). Regardless of these ambiguities, it is clear from Figure 5 that the Sotho languages started to be increasingly discussed from 1997 onwards, and peaked for about seven years, from 2001 until 2007. The main language within the family being discussed is Northern Sotho (with often, for *Lexikos* 12 for example, all research articles devoted to it).

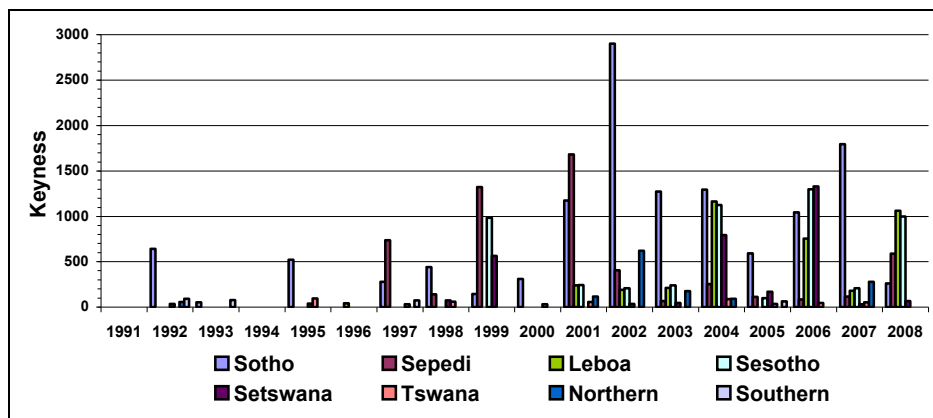


Figure 5: Keyness values for 'Sotho languages' in each *Lexikos* issue

Moving to the Nguni family of southern Bantu languages, one notes that one member is ambiguous: 'Ndebele', as it can refer to the Ndebele spoken in South Africa (isiNdebele in the language) or Zimbabwean Ndebele. The extreme values for Ndebele in Figure 6 are actually mostly for Zimbabwean Ndebele.

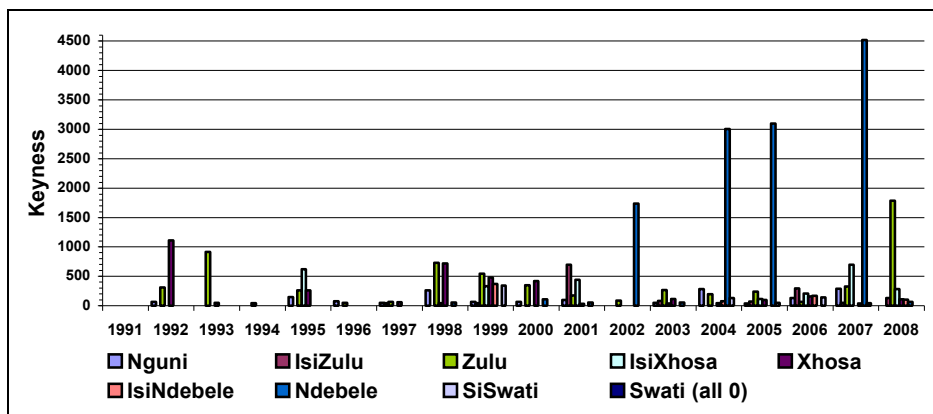


Figure 6: Keyness values for 'Nguni languages' in each *Lexikos* issue

Compared to the Sotho family, then, the Nguni family of languages is not frequently discussed in *Lexikos*, except for the period 1998–2001, and starting with 2008 for Zulu. 1992 was also a good year for Xhosa.

To complete the picture for the official South African languages, Figure 7 displays the keyness values for Venda (Tshivenda) and Tsonga (Xitsonga). Both these languages are not often discussed, with their keyness values on the whole (much) lower than for the other official South African languages.

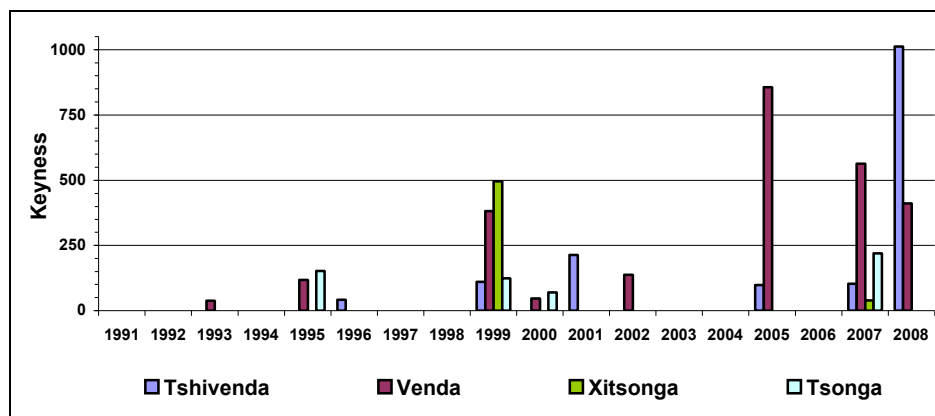


Figure 7: Keyness values for *Venda* and *Tsonga* in each *Lexikos* issue

3.4 Key topics in *Lexikos*

Most of the remaining keywords in Addenda 1 to 3 are true keywords. Analysing those brings out the real features of a journal like *Lexikos*. A few groups of related keywords and their patterns will be looked into below and, where relevant, comparisons will be made with the corresponding patterns found in IJL (cf. De Schryver 2009a).

3.4.1 Parts of Speech (POSS)

In Table 16 the parts of speech (POSS) as well as a few related terms relevant for the study of Bantu languages — all excerpted from the addenda — have been tabulated. The differences between the three phases are striking. In Phase 1, only two POSS stand out: *verbs* and *nouns* (together with the related *plural*), and for the Bantu languages also *stem*, *reflexive*, *ideophone*, and *prefix*. Going to Phase 2, *copulative* (and also *singular*) is added to the basic core; and going to Phase 3, *adverb*, *demonstrative*, *adjective*, and *pronoun* are further added, as well as *suffix* and *cl.* (class) for the Bantu languages, and even *POS* (part of speech) itself. The list of word classes studied from a lexicographic point of view most definitely grows with the years in *Lexikos*, and in that it is the opposite from the

IJL pattern. Actually, following the IJL study we 'predicted' that such lists of POSs would only grow whenever 'new' languages would become the object of lexicographic investigation. This is the case in *Lexikos*, where many of the languages are being described for the very first time.

Table 16: Parts of speech (POSs) in the three *Lexikos* sub-corpora

Phase 1 (1991–1996)		Phase 2 (1997–2002)		Phase 3 (2003–2008)	
Keyword	Keyness	Keyword	Keyness	Keyword	Keyness
VERB	2440.16	VERB	2335.04	VERB	1791.75
VERBS	1782.64	NOUN	2218.57	NOUN	1481.39
NOUN	1305.44	NOUNS	1327.65	ADVERBS	1314.56
NOUNS	1209.48	COPULATIVE	1091.27	NOUNS	1219.52
STEM	596.12	VERBS	1054.38	VERBS	1191.18
REFLEXIVES	562.76	COPULATIVES	798.12	DEM	1068.77
IDEOPHONES	450.20	PLURAL	734.68	ADJ	981.19
PLURAL	436.76	SINGULAR	609.03	SUFFIX	980.06
PREFIX	387.42	PREFIX	549.19	ADJECTIVES	912.92
PREFIXES	287.91	STEM	487.98	ADJECTIVE	870.45
		PLURALS	440.21	CL	830.74
				PRON	820.91
				POS	748.93
				PREFIX	665.35
				COPULATIVE	655.69
				VERBAL	651.68

3.4.2 N-lingual lexicography

Monolingual, bilingual, trilingual and multilingual lexicography are all being discussed in *Lexikos*, but not in equal proportions, and those proportions change with time, as summarized in Figure 8.

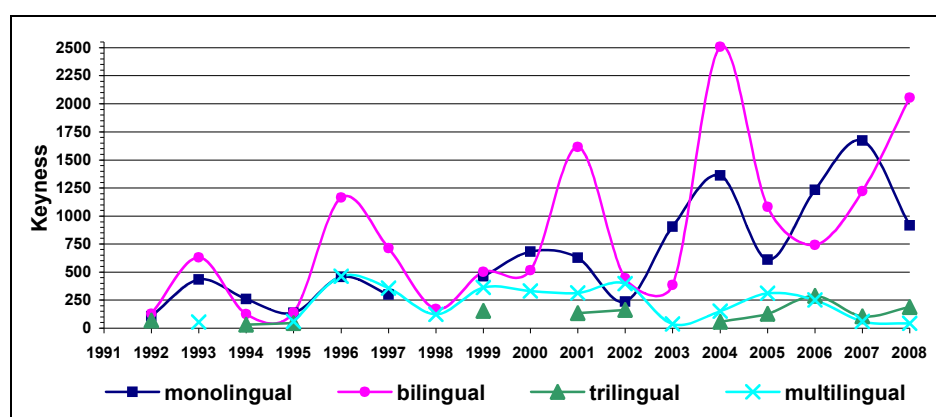


Figure 8: Keyness values for n-lingual in each *Lexikos* issue

Overall, bilingual lexicography is more popular than monolingual and multilingual lexicography in lexicographic discussions. While trilingual lexicography only appears sporadically, multilingual topics are actually slightly decreasing. Both bilingual and monolingual lexicography are on the rise, with the acceleration of the former greater than that of the latter.

The core task in monolingual lexicography is defining, so one could investigate whether the occurrence of key terms like *defining*, *definition(s)* or *explanatory* have changed during the past eighteen years. Apart from one peak for *definitions* in 2007, all these key terms remained stable, as seen in Figure 9.

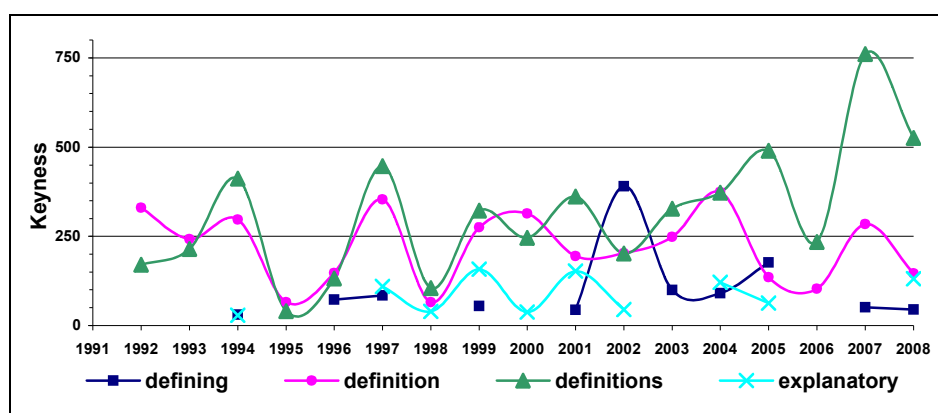


Figure 9: Keyness values for *defining* etc. in each *Lexikos* issue

Likewise, the core task in bilingual lexicography is to provide translation equivalents, so Figure 10 shows whether or not changes have occurred in this regard, looking at the key terms *equivalence*, *equivalent(s)*, *translating*, and *translation(s)*. Here one notices an increased activity for the period 2002–2004.

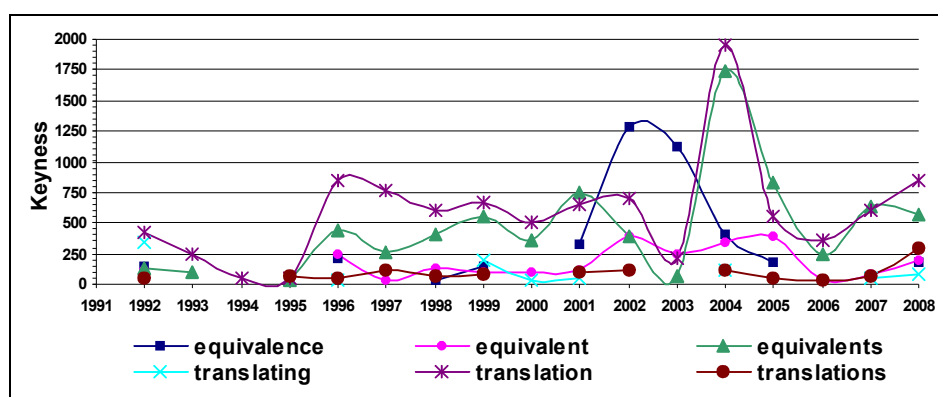


Figure 10: Keyness values for *translating* etc. in each *Lexikos* issue

3.4.3 Terminology and terminography

With regard to the sub-disciplines of terminology and terminography, Figure 11 clearly indicates that there has been a flurry of activity in the period 1997–2004, especially in the years 2000 and 2002.

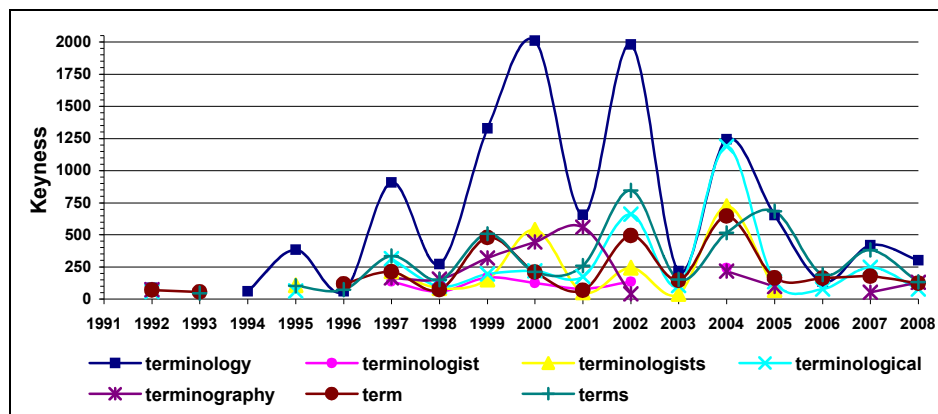


Figure 11: Keyness values for *terminology etc.* in each *Lexikos* issue

3.4.4 Metalexicography

In Section 3.2 above, we noted *Lexikos*'s openness towards metalexicography. We are now in a position to analyse this aspect in more depth. Figure 12 plots the distribution of the general keywords *concept(s)*, *principle(s)*, *scientific* and *theoretical*. From the graph it is clear that the period 1999–2004 should be regarded as the most active. Focusing on lexicography itself, this is confirmed in Figures 13 through 15.

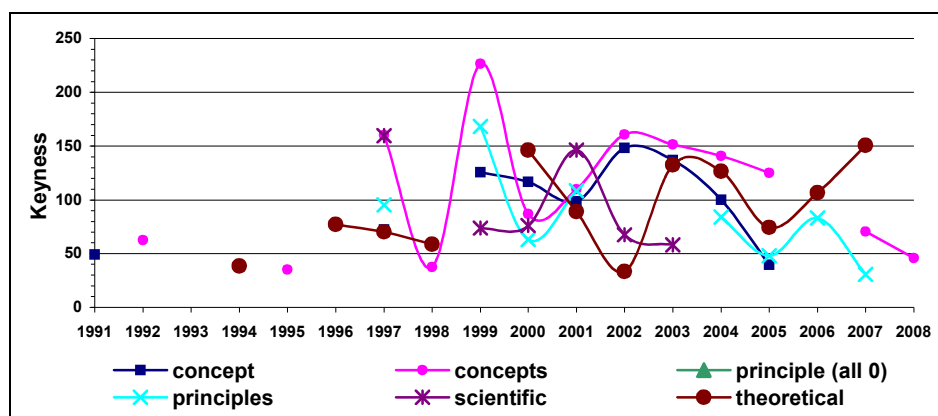


Figure 12: Keyness values for *theoretical etc.* in each *Lexikos* issue

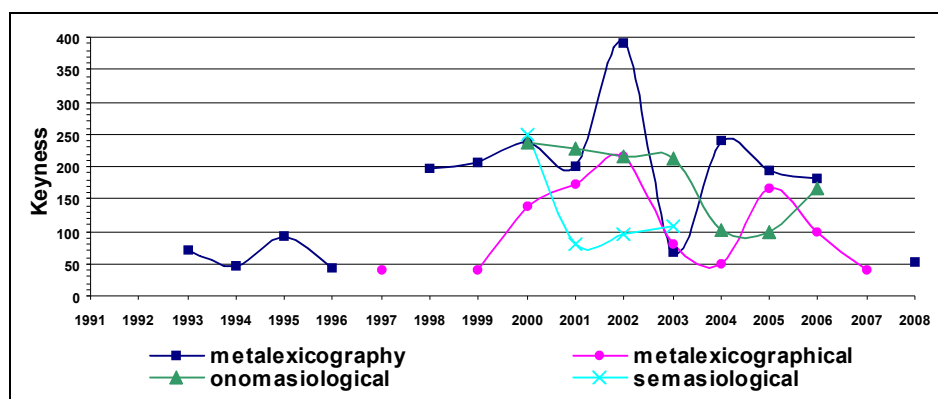


Figure 13: Keyness values for *metalexigraphy* etc. in each *Lexikos* issue

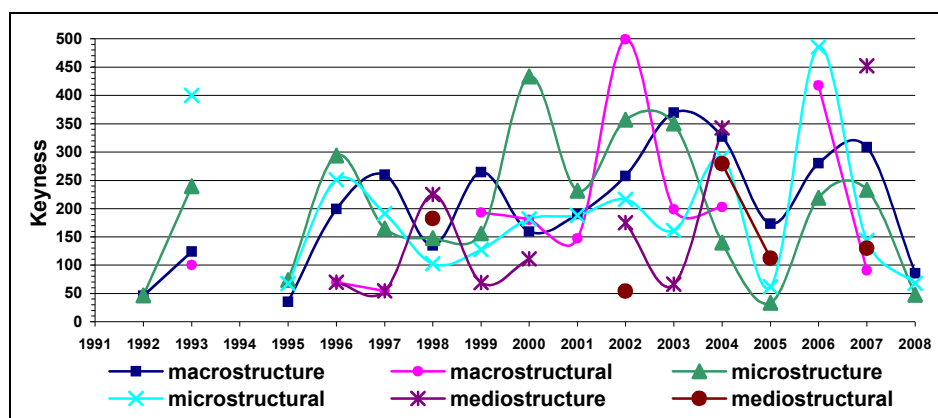


Figure 14: Keyness values for *x-structure* in each *Lexikos* issue

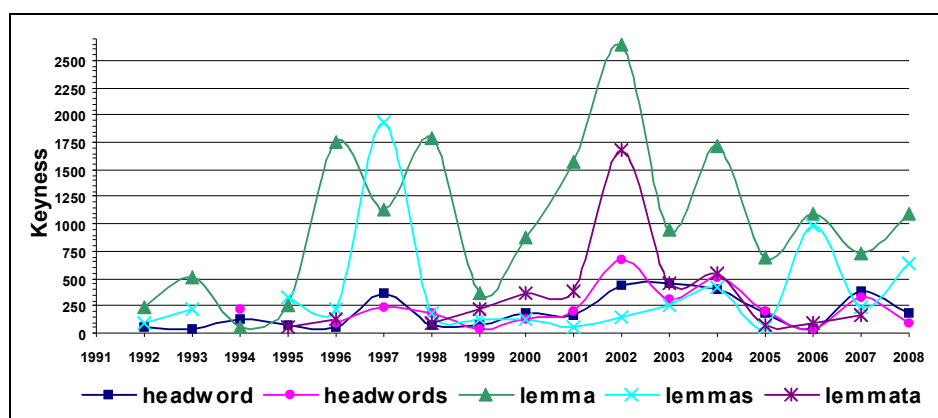


Figure 15: Keyness values for *lemma* etc. in each *Lexikos* issue

In Figure 13, the terms *metalexicography* and *metalexicographical* are especially key for the period 2000–2005, while highly technical terms like *onomasiological* vs. *semasiological* also become key during that period. Figure 14 looks at where terms like *macro-*, *micro-* and *mediostucture* peak, which is roughly 2000–2004. In Figure 15, lastly, the common *headword(s)* is contrasted with the scientific *lemma(s)/lemmata*, which clearly indicates that the vast majority of the contributors to *Lexikos* indeed use field-specific terms.⁷ These results stand in sharp contrast to those found for IJL, where far fewer theoretical terms are found in the keyword lists.

3.4.5 Polysemy and synonymy

Figures 16 and 17 summarize the use of respectively polysemous and synonymous terms in *Lexikos*.

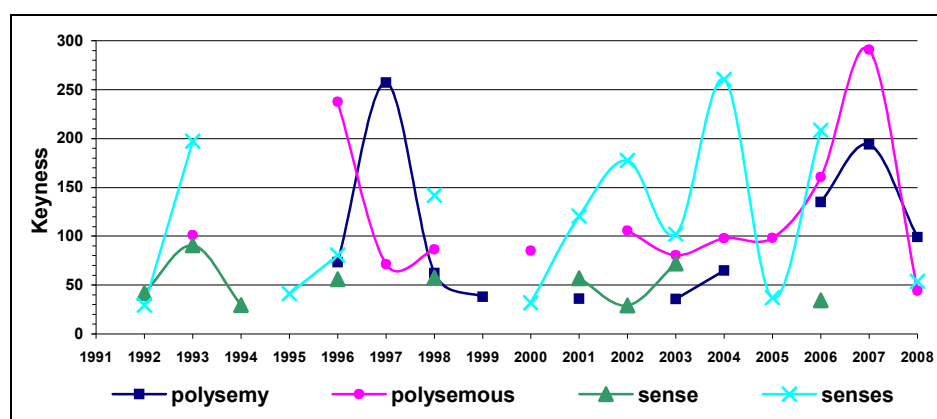


Figure 16: Keyness values for *polysemy* etc. in each *Lexikos* issue

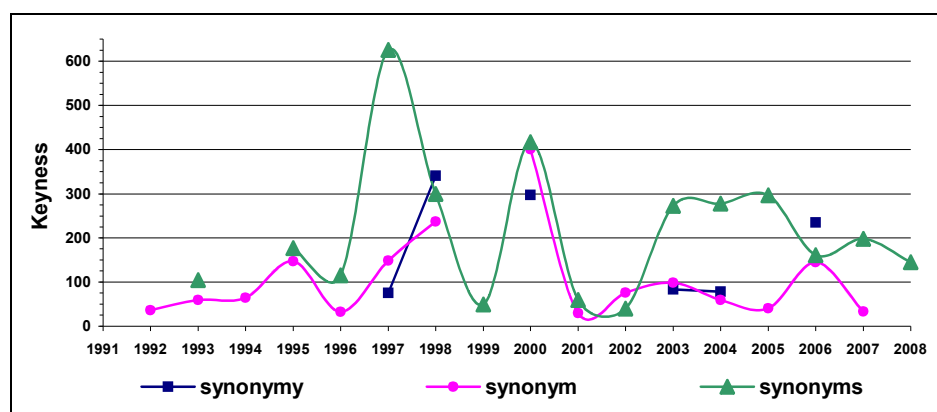


Figure 17: Keyness values for *synonymy* etc. in each *Lexikos* issue

From Figures 16 and 17 one may conclude that polysemy and synonymy are (a) not typically treated together in metalexicographical discussions, and (b) not always at the centre of such discussions. Both these findings run against what one would have expected.

3.4.6 Meaning, collocations, and examples vs. idioms/proverbs, and dialects/slang

While there are lexicographic topics for which one expects some kind of continuity, such as *meaning*, *collocations* or *examples*, others only appear sporadically, such as *idioms* and *proverbs* or *dialects* and *slang*. The difference can clearly be seen from Figures 18 through 22. Such graphs thus visualize where which topics are being discussed, and with which weighted frequency. In each of those graphs, peaks nonetheless pinpoint heightened activity.

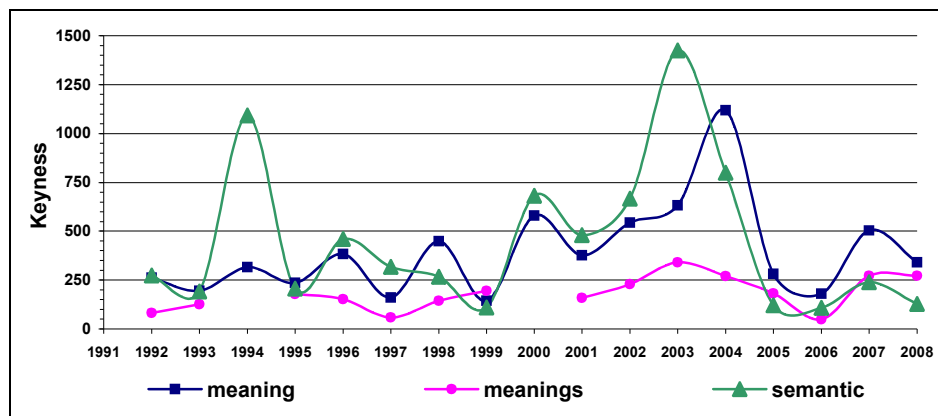


Figure 18: Keyness values for *meaning(s)* and *semantic* in each *Lexikos* issue

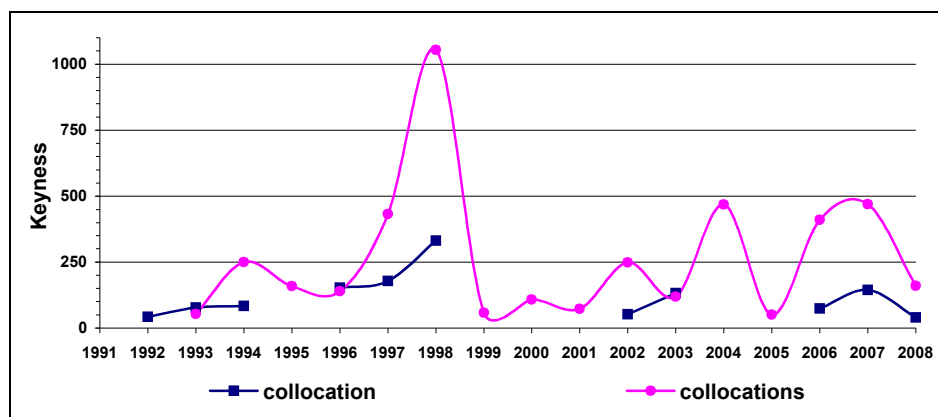


Figure 19: Keyness values for *collocation(s)* in each *Lexikos* issue

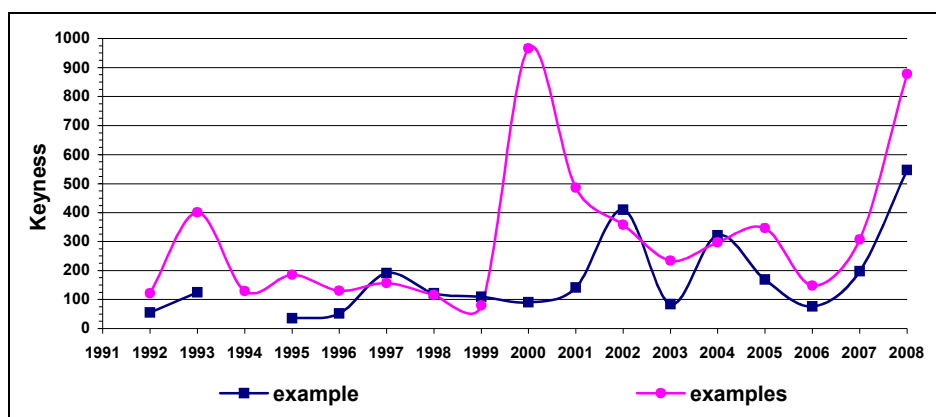


Figure 20: Keyness values for *example(s)* in each *Lexikos* issue

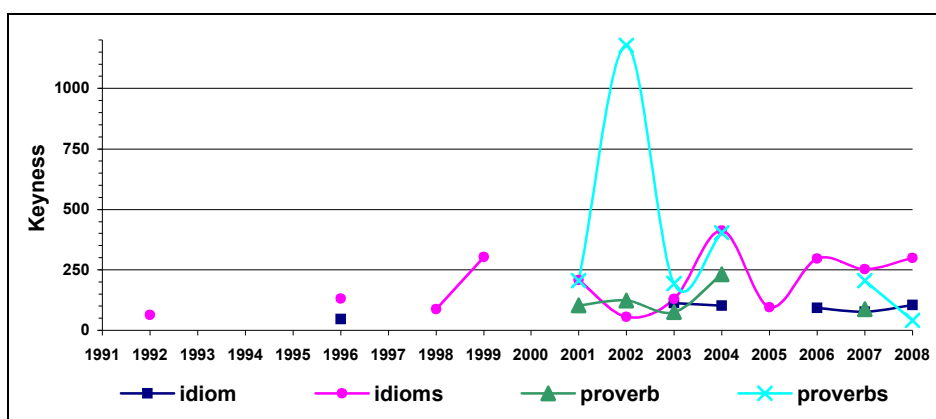


Figure 21: Keyness values for *idiom(s)* and *proverb(s)* in each *Lexikos* issue

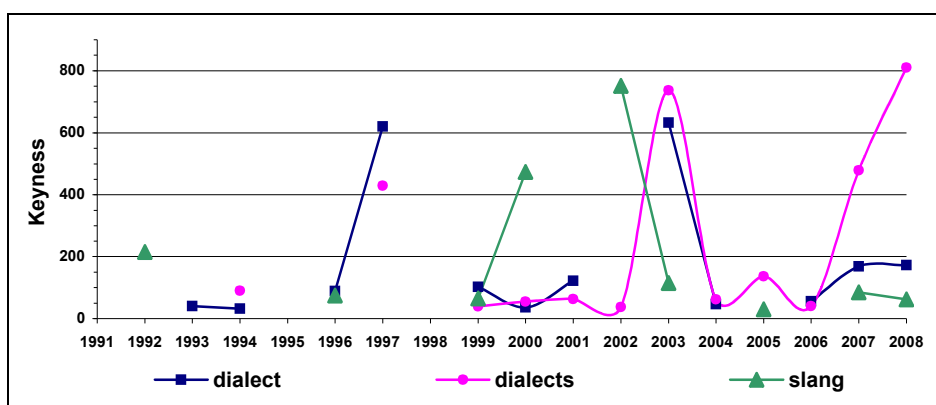


Figure 22: Keyness values for *dialect(s)* and *slang* in each *Lexikos* issue

3.4.7 The user perspective

To conclude this section, one may also briefly look at the user perspective. While the high day of the user perspective was found to be the period 1997–2004 in IJL, Figures 23 and 24 indicate that that period is narrowed down in *Lexikos* to basically two years: 2001 and 2004 (recall the special section on 'Learners' Dictionaries' that was added in Volume 14, cf. Section 2.1 above).

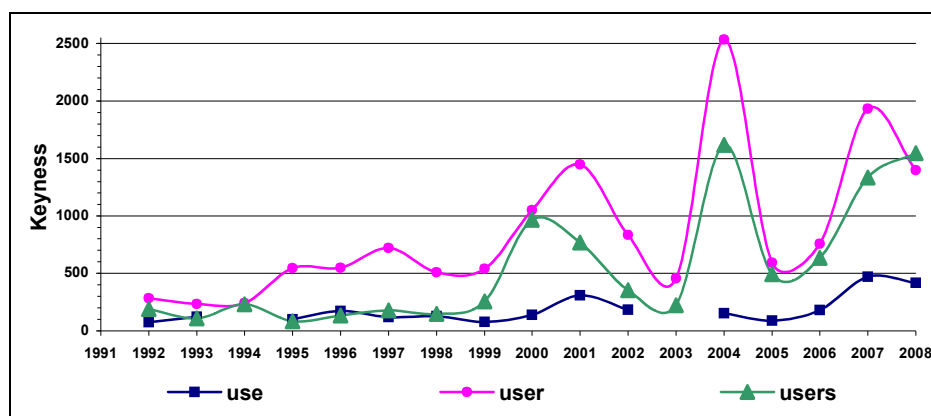


Figure 23: Keyness values for *use* and *user(s)* in each *Lexikos* issues

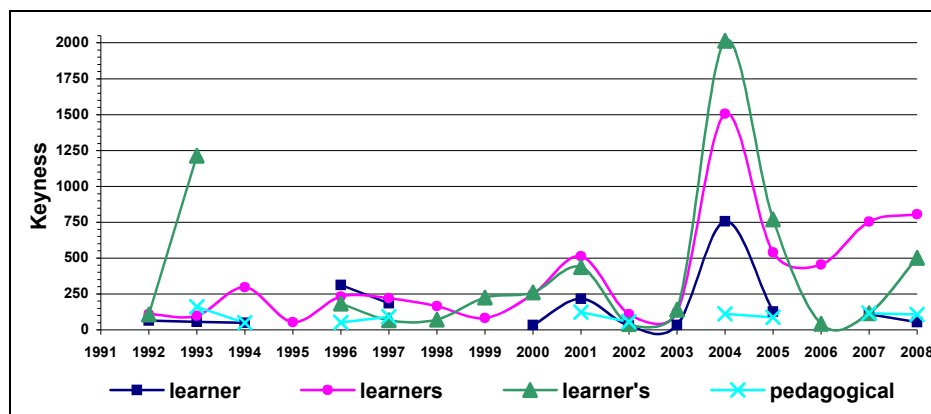


Figure 24: Keyness values for *learner(s)* etc. in each *Lexikos* issue

4. Summary and discussion

In this article an in-depth analysis was presented of both the publication details and the characteristic lexical items for the journal *Lexikos* from the first volume to the present. In Section 1 it was pointed out that *Lexikos* has become increasingly more professional with each new volume. Section 2 was devoted to an

analysis of the *Lexikos* Metadata Database. It was shown that the number of pages as well as the number of contributions — averaging 366 pages and 30 contributions per volume — steadily rose over the years. English has become the main discourse language (currently stabilized at 70%, on average 65%), followed by Afrikaans (currently stabilized at 20%, on average 26%). Other languages used are German, French, and Dutch. About 44% of the contributions are articles, 4% are review articles, and 20% are book reviews. As many as 83% of the contributions are single-authored, 12% are joint efforts. A total of 260 scholars contributed to *Lexikos* during the first eighteen years, one third of them (34%) more than once. Those repeat contributors are also 'faithful' in that their input is regular. Most repeat contributors moreover started publishing in *Lexikos* over the past few years, which bodes well for the future of *Lexikos*. About three quarters of the contributors (74%) are based in Africa, 86% of those in South Africa. Within South Africa, 30% are based in Pretoria and as many as 46% in Stellenbosch. In Pretoria, the most productive centre is the Department of African Languages at the University of Pretoria (33%), while in Stellenbosch it is the Department of Afrikaans and Dutch at the University of Stellenbosch (39%) and especially the Bureau of the WAT (45%).

Some caution in interpreting the last set of results is in order. While it is undeniable that a very large percentage of the material published in *Lexikos* emanates from South Africa, the African reach and especially the interaction with the rest of Africa is larger than these statistics suggest. Firstly, through various cooperation agreements, students from the rest of Africa are enrolled at South African universities, where they obtain degrees in lexicography. The prime example here is the Department of Afrikaans and Dutch at the University of Stellenbosch, where several Gabonese students have obtained their PhDs over the past few years. Many of them published their results in *Lexikos*, using their South African affiliations while in South Africa. This is for example the case for Mavoungou, nr. 12 in Table 6. The same is true for PhD students from Namibia (for example Beyer, nr. 16 in Table 6). Secondly, success breeds success. Both *Lexikos* and AFRILEX can definitely be seen as South African lexicographic success stories, and as a result of their reputation, scholars now continue to visit South Africa, wishing to do their research in South Africa rather, sometimes even on South African languages (as is the case for the current writer).

Section 3 was devoted to an analysis of the *Lexikos* Corpus, more in particular that part containing all 355 English-language contributions which appeared in *Lexikos* to date, totalling 1 744 536 tokens and 65 485 types. That corpus was partitioned in three sub-corpora (of 6 years each) as well as 18 sub-corpora (i.e. per year), in order to study various trends across time (based on the keyness measure as compared with the BNC). With regard to dictionaries discussed in *Lexikos*, a rather stable picture was found, both in terms of number of keywords and keyness values. In contrast, the number of lexicographers competing for keyword status increases with time. Looking at the languages as

objects of investigation, those were found to be overwhelmingly English and Afrikaans first, and biased towards the official South African languages second. Within the latter group, the order of attention is (a) Sotho languages (with especially Northern Sotho), (b) Nguni languages (with Zulu the most recent focus), (c) Venda, and lastly (d) Tsonga.

Seven key topic areas were also studied, the first being parts of speech, for which it was found that the list of word classes studied from a lexicographic point of view most definitely grows with the years in *Lexikos*. The second area was *n*-lingual lexicography, for which it was shown that bilingual lexicography is more popular than monolingual and multilingual lexicography in scholarly discussions. Thirdly, when it comes to terminology and terminography, a flurry of activity for the period 1997–2004 could be pinpointed. In a fourth section *Lexikos* was found to be highly theoretically oriented, certainly more than the *International Journal of Lexicography*, with the vast majority of the contributors using field-specific terms. Fifthly, for polysemy and synonymy the conclusions were that they are (a) not typically treated together in metalexicographical discussions, and (b) not always at the centre of such discussions. In a sixth section continuity was found for topics such as meaning, collocations or examples, while topics such as idioms and proverbs or dialects and slang are more sporadic. Lastly, attention to the user perspective in *Lexikos* was found to peak in the years 2001 and 2004.

Looking back, then, one cannot but be impressed by the achievements of *Lexikos*. One recognizes a healthy research culture, to which both lexicographers and metalexicographers contribute, on a variety of languages, in a variety of sub-disciplines. There is still a strong South African bias, however, which *Lexikos* as an adult should now shed.

Acknowledgements

Many thanks are due to Ms Riette Ruthven, who helped in assembling the *Lexikos* 1–10 material underlying the *Lexikos* Corpus. She managed to extract most data from legacy Word files, re-scanned the missing parts, and helped restoring diacritics. Volumes 11–17 were downloaded via the Sabinet Online service. *Lexikos* 18 was received directly as a single PDF file from the Bureau of the WAT.

Endnotes

1. Only very rarely is an article a reworked version of a published conference proceedings paper, such as Gouws (2001) vs. Gouws (2001a), or Mafela (2004) vs. Mafela (2005).
2. The percentage of editorial material is relatively high. This is so because all editorial material — written by the *Lexikos* Editor and the AFRILEX President — always appears in both Afrikaans and English.

3. In calculating this, some of the names were collapsed: 'Potgieter, Liezl' with (née) 'Gouws, Liezl', and 'Nong, Salmina' with (née) 'Maepa, Salmina'.
4. See endnote 2.
5. For the calculation of keywords and associated keyness values, the log-likelihood statistic of WordSmith 4.0 was used, with the minimum keyword frequency set at 3, and the maximum p value = 0.0000001. See for more details on the extraction of keywords for lexicographic purposes, Taljard and De Schryver (2002).
6. As keyness values are sensitive to (sub-)corpus size, figures for six-year windows are not directly comparable to per-year figures.
7. As one anonymous adjudicator correctly pointed out, this statement assumes (since these are keyness values, not frequencies) that the lemmas [headword] and [lemma] are of equal frequency in general English / the BNC. This is indeed a risky assumption, but as it turns out, is true: In the lemmatized BNC, the occurrences are respectively 65 vs. 60.

References

- BNC World Edition. 2009. A 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century [online]. <http://www.natcorp.ox.ac.uk/>
- De Schryver, G.-M. 1999. *Bantu Lexicography and the Concept of Simultaneous Feedback, Some Preliminary Observations on the Introduction of a New Methodology for the Compilation of Dictionaries with Special Reference to a Bilingual Learner's Dictionary Cilubà-Dutch*. Unpublished M.A. Dissertation. Ghent: Ghent University.
- De Schryver, G.-M. 2009. A Few Words from AFRILEX. *Lexikos* 19: xiv-xv.
- De Schryver, G.-M. 2009a. Bibliometrics in Lexicography. *International Journal of Lexicography* 22(4).
- Du Plessis, J.C.M.D. 2008. Foreword. *Lexikos* 18: xi.
- Gouws, R.H. 2001. Enhancing the Access Structure of Translation Dictionaries. Lee, S. (Ed.). 2001. *ASIALEX 2001 Proceedings, Asian Bilingualism and the Dictionary*: 77-83. Seoul: Yonsei University, Center for Linguistic Informatics Development.
- Gouws, R.H. 2001a. The Use of an Improved Access Structure in Dictionaries. *Lexikos* 11: 101-111.
- Harteveld, P. 1991. Foreword. *Lexikos* 1: xii-xiv.
- Mafela, M.J. 2004. Bilingual Dictionary and Meaning Discrimination in Venda. Williams, G. and S. Vessier (Eds.). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6-10, 2004*: 501-509. Lorient: Université de Bretagne Sud, Faculté des Lettres et des Sciences Humaines.
- Mafela, M.J. 2005. Meaning Discrimination in Bilingual Venda Dictionaries. *Lexikos* 15: 276-285.
- Scott, M. 2009. *WordSmith Tools* [online]. <http://www.lexically.net/wordsmith/>
- Taljard, E. and G.-M. de Schryver. 2002. Semi-automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho. *Lexikos* 12: 44-74.
- Web of Science. 2009. Provides researchers, administrators, faculty, and students with quick, powerful access to the world's leading citation databases. Authoritative, multidisciplinary content covers over 10 000 of the highest impact journals worldwide [online]. <http://apps.isiknowledge.com/> [subscription required]
- Zgusta, L. 1993. Book Review: Harteveld, P., Ed. 1991. *Lexikos. International Journal of Lexicography* 6(4): 305-306.

Addendum 1: Keywords and keyness values for *Lexikos* sub-corpus 1: 1991–1996

#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness
1	DICTIONARY	20062.88	76	Y	874.20	151	BANTU	441.47	226	STANDARD	298.10
2	AFRIKAANS	11942.42	77	X'S	856.70	152	SEMANTICS	440.44	227	ARTICLES	295.97
3	DICTIONARIES	9584.93	78	ENTRY	846.61	153	METAFORMS	438.95	228	MALAYO	292.63
4	ENGLISH	8755.20	79	ITEM	832.12	154	WÖRTERBUCH	438.95	229	WOORDENBOEK	292.63
5	LANGUAGE	6035.01	80	BENG	826.54	155	PLURAL	436.76	231	LIT	292.02
6	LEXICOGRAPHY	5827.79	81	TEXTUAL	824.81	156	CITATIONS	433.80	230	DEROGATORY	289.50
7	LEXICAL	4967.50	82	TEXT	823.28	157	ET	430.71	232	MORPHEME	288.87
8	AFRICAN	4861.51	83	MALAYU	810.37	158	ALLEX	427.69	233	OFFENSIVE	287.93
9	LANGUAGES	4281.88	84	NIE	808.10	159	NEUTRAL	425.71	234	PREFIXES	287.91
10	DUTCH	3783.88	85	EDS	781.55	160	TL	423.08	235	TERM	287.78
11	LEXEMES	3415.93	86	MAGN	780.80	161	REFERENCES	420.90	236	YA	287.69
12	WORD	3198.02	87	HAUSMANN	766.13	162	STELLENBOSCH	417.49	237	MORPHEMES	286.96
13	WAT	3039.17	88	ARTICLE	752.27	163	NATIVE	416.80	238	STRUCTURE	286.71
14	SOUTH	2817.94	89	LEARNERS	744.21	164	OXFORD	411.96	239	AFRIKA	286.63
15	VERB	2440.16	90	LEMMAS	734.55	166	ENCYCLOPEDIA	409.48	240	TRANSLATING	284.49
16	USAGE	2435.08	91	USERS	721.04	165	SYNONYMS	405.81	241	FURCHT	281.38
17	LEXICOGRAPHICAL	2427.54	92	NACC	694.09	168	KLA	405.18	242	GROLL	281.38
18	LEMMA	2413.09	93	AMERICAN	692.96	169	LEMMATIC	405.18	243	KAMUSI	281.38
19	LEXICOGRAPHER	2357.89	94	BLOUNT	686.81	167	ÜBERWINDEN	405.18	244	VERFLIEGEN	281.38
20	SWAHILI	2202.69	95	AFK	686.57	170	PRONUNCIATION	400.25	245	EQUIVALENCE	281.21
21	RACIST	2112.91	96	AFRILEX	686.57	171	GRAPHS	395.77	246	MORPHOLOGICAL	280.59
22	SEMANTIC	2043.33	97	REFERENCE	661.17	172	BÖSCHUNG	393.93	247	ABBREVIATIONS	280.26
23	NON	2002.94	98	FR	648.36	173	IPA	390.51	248	SLANG	280.11
24	BILINGUAL	1998.65	99	EQUIVALENTS	640.55	174	IN	390.49	249	PHONETIC	277.58
25	WORDS	1919.72	100	LINGUISTS	635.97	175	DB	389.23	250	EXAMPLE	277.41
26	USER	1799.97	101	OCCURRENCE	627.08	176	TEXTS	389.17	251	NIGGER	274.58
27	LEXICOGRAPHERS	1791.22	102	CONDENSATION	624.27	177	PREFIX	387.42	252	DISCUSSION	273.04
28	VERBS	1782.64	103	EDITORIAL	611.66	178	ASHKENAZIC	384.80	253	DSYNTA	270.12
29	LINGUISTIC	1769.81	104	USAGES	600.86	179	LEXICOGRAPHICA	382.67	254	LEXIKOGRAFIE	270.12
30	YIDDISH	1664.82	105	IS	600.30	180	WUT	382.67	255	OSSELTON	270.12
31	ENTRIES	1644.69	106	STEM	596.12	181	FORM	378.27	256	SCHÜREN	270.12
32	LEXICOGRAPHIC	1638.07	107	MICROSTRUCTURAL	590.54	182	EQUIVALENT	378.20	257	DICTIONNAIRE	270.04
33	CF	1633.09	108	MICROSTRUCTURE	577.50	183	ORTHOGRAPHIC	374.30	258	TONO	269.36
34	AFRICA	1628.05	109	ANGST	564.87	184	AL	367.56	259	OED	267.87
35	ETYMOLOGY	1525.77	110	REFLEXIVES	562.76	185	ILLUSTRATIVE	367.24	260	CATEGORIES	267.04
36	TRANSLATION	1489.92	111	BRANFORD	559.64	186	WNT	360.16	261	ELKE	266.36
37	ITEMS	1452.08	112	UNIVERSITY	550.47	187	SHONA	358.27	262	HABEN	265.28
38	WOORDEBOEK	1384.41	113	LFS	549.51	188	HOFFNUNG	357.09	263	PRESS	264.49
39	ZULU	1365.02	114	TERMINOLOGY	547.88	189	MACROSTRUCTURE	352.84	264	AUSTRALIAN	261.75
40	MEANING	1359.95	115	CORPORA	547.11	190	WÖRTERBÜCHER	348.91	265	PEDAGOGICAL	260.47
41	CAPE	1336.03	116	MULTILINGUAL	537.95	191	ALPHABET	344.36	266	INDIGENOUS	260.43
42	JEWISH	1310.92	117	MEANINGS	537.46	192	POLYSEMOUS	343.40	267	ONESELF	260.39
43	NOUN	1305.44	118	LF	529.46	193	CONCISE	342.38	268	ÄRGER	258.87
44	HARTMANN	1301.68	119	ORTHOGRAPHIES	523.47	194	LABEL	340.78	269	DICTIONNAIRES	258.87
45	CORPUS	1266.75	120	COLLOCATIONS	523.05	195	KISI	337.65	270	EMPFINDEN	258.87
46	LEARNER'S	1258.91	121	SPELLING	513.43	196	GRAMMAR	336.41	271	ENTZÜCKEN	258.87
47	MONOLINGUAL	1229.53	122	ORIGIN	513.31	197	EXPRESSIONS	333.49	272	ERREGEN	258.87
48	NOUNS	1209.48	123	USED	509.56	198	FORMS	331.60	273	ILSON	258.87
49	TAAL	1188.01	124	OF	506.45	199	NORTHEASTERN	331.40	274	JOOD	258.87
50	AFRIKAANSE	1181.81	125	PRETORIA	498.75	200	COBUILD	330.31	275	MANIFESTABLE	258.87
51	LINGUISTICS	1173.62	126	ZGUSTA	495.23	201	NIEMEYER	328.64	276	TWEETALIGE	258.87
52	COMPILERS	1169.18	127	LEXICOGRAPHER'S	495.19	202	HEADWORD	328.39	277	UNTERDRÜCKEN	258.87
53	LEXEME	1165.18	128	EDITION	491.29	203	BEGEISTERUNG	326.40	278	VERSETZEN	258.87
54	EMOTION	1162.40	129	SYNTACTIC	483.07	204	PANIK	326.40	279	WECKEN	258.87
55	GRAMMATICAL	1154.34	130	KISWAHILI	480.59	205	COLLOCATION	325.44	280	VOWEL	257.64
56	ARABIC	1151.52	131	ETC	479.51	206	SENSES	323.63	281	SUCH	256.82
57	BUREAU	1149.35	132	HOTTENTOT	478.79	207	NASAL	323.29	282	DATA	255.94
58	XHOSA	1144.14	133	NADEN	478.79	208	INCLUSION	321.69	283	TERMS	252.47
59	ACC	1090.04	134	MUSLIM	478.49	209	FEM	317.87	284	VS	250.87
60	LEXIKOS	1080.51	135	LEXICON	473.80	210	EIFERSUCHT	315.14	285	LEGEN	247.61
61	SPEAKERS	1071.30	136	GOUWS	472.71	211	FÜHLEN	315.14	286	MEL'CUK	247.61
62	NETHERLAND	1057.53	137	ISIXHOSA	472.71	212	GIVEN	313.79	287	NACHLASSEN	247.61
63	DEFINITION	1035.40	138	USE	471.46	213	SICH	312.65	288	OTJHERERO	247.61
64	GEFÜHL	1012.97	139	ACHTUNG	471.20	214	SYNONYM	311.47	289	SCHAM	247.61
65	ORTHOGRAPHY	985.06	140	FREUDE	467.72	215	LOB	310.39	290	ÜBERKOMMEN	247.61
66	WIEGAND	984.95	141	BL	464.54	216	THUS	310.28	291	LIEBE	244.88
67	ED	952.68	142	TÜBINGEN	461.46	217	TARGET	306.93	292	HEADWORDS	243.54
68	VOCABULARY	947.04	143	NTS	459.10	218	BEKOMMEN	303.89	293	CONDENSED	242.32
69	EXAMPLES	946.51	144	LINGUIST	454.15	219	EURALEX	303.89	294	BRUIN	240.94
70	SOTHO	945.81	145	LEARNER	453.59	220	LEIDENSCHAFT	303.89	295	GHANA	239.69
71	INFORMATION	944.23	146	IDEOPHONES	450.20	221	OR	303.39	296	NGUNI	239.39
72	DEFINITIONS	922.19	148	INL	450.20	222	COMPILATION	302.90	297	CATEGORY	237.46
73	KAFFIR	921.47	147	KAFFER	450.20	223	ECD	300.46	298	LABELS	236.93
74	NDAT	900.42	149	NAMES	443.92	224	ISO	300.29	299	BEDAVERN	236.36
75	PHRASAL	896.81	150	EXETER	443.51	225	AS	299.12	300	MITLEID	236.36

Addendum 2: Keywords and keyness values for *Lexikos* sub-corpus 2: 1997–2002

#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness
1	DICTIONARY	37222.13	76	ITEMS	1536.58	151	ELECTRONIC	783.27	226	DEFINING	538.12
2	DICTIONARIES	19532.26	77	BOHLALE	1525.54	152	MICROSTRUCTURAL	779.39	227	CONCEPT	537.93
3	LANGUAGE	13562.42	78	EQUIVALENCE	1521.71	153	MEANINGS	764.71	228	LIST	535.53
4	LEXICOGRAPHY	11015.13	79	AFRIKAANSE	1515.44	154	AL	757.99	229	PANSALB	535.45
5	LANGUAGES	8885.86	80	NDEBELE	1485.13	155	REFERENCES	756.41	231	GABONESE	530.05
6	DIE	8761.30	81	GRAMMATICAL	1455.08	156	KNOWLEDGE	750.00	230	SEMI	528.63
7	AFRICAN	7485.60	82	TERM	1410.82	157	TONGUE	749.67	232	SMIT	525.56
8	CORPUS	7412.14	83	WOORDEBOEKE	1384.10	158	CONCEPTS	748.61	233	PHRASES	525.15
9	WIEGAND	7095.03	84	TERMINOLOGICAL	1374.35	159	LOANWORDS	739.18	234	DRC	521.89
10	ENGLISH	6987.90	85	WIEGAND'S	1353.79	160	GABON	738.47	235	TEXTUAL	518.29
11	VAN	6791.82	86	DEFINITION	1347.76	161	SUID	737.50	236	DUTCH	519.55
12	LEMMA	6682.36	87	UNIVERSITY	1344.69	162	PLURAL	734.68	237	SETSWANA	518.29
13	TERMINOLOGY	6400.04	88	XHOSA	1335.53	163	FIELD	725.42	238	NOORD	515.43
14	WORD	6174.76	89	NOUNS	1327.65	164	DIALECT	723.68	239	RECHISHONA	515.24
15	LEXICAL	5329.38	90	MPŠA	1293.17	166	STANDARD	721.04	240	TERMINOLOGIST	515.24
16	AFRIKAANS	5313.14	91	SWAHILI	1289.27	165	GRUYTER	716.85	241	PARAPHRASE	514.66
17	LEXICOGRAPHIC	5216.82	92	SYNONYMS	1287.77	168	MOSADI	707.19	242	CATEGORIES	510.05
18	USER	4914.39	93	SCHRYVER	1261.29	169	DICTIONNAIRE	706.04	243	EXPLANATORY	509.26
19	WORDS	4630.34	94	LEARNERS	1247.80	167	FORMS	701.07	244	LEKSIKOGRAFIESE	505.14
20	SOTHO	3946.56	95	MICROSTRUCTURE	1242.85	170	CB	699.17	245	ONOMASIOLOGICAL	505.14
21	TRANSLATION	3693.35	96	TERMINOGRAPHY	1242.65	171	ENCARTA	697.09	246	LINGUISTS	498.46
22	BILINGUAL	3542.47	97	ET	1242.16	172	LEMMATIZATION	697.09	247	SPELLING	497.15
23	HAUSA	3424.94	98	ALPHABETICAL	1237.81	173	TÜBINGEN	697.09	248	NAMELY	495.55
24	WAT	3318.27	99	ORTHOGRAPHY	1208.73	174	INTERNET	686.18	249	WOORDE	495.03
25	SEPEDI	3212.81	100	HEADWORDS	1199.94	175	ALBERTS	682.21	250	WÖRTERBÜCHER	495.03
26	LINGUISTIC	3160.48	101	NIE	1164.97	176	NA	680.52	251	MACQUARIE	492.82
27	ARTICLE	3079.72	102	BA	1140.54	177	CONSIDERATE	677.78	252	EDITION	492.50
28	SHONA	3058.44	103	TAAL	1123.86	178	COMPILED	676.60	253	SCIENTIFIC	490.98
29	TEXT	2955.75	104	SESOTHO	1122.08	179	SOURCE	673.48	254	FUNCTIONAL	490.14
30	CF	2710.90	105	PROVERBS	1106.09	180	DIPUKU	666.78	255	STEM	487.98
31	WOORDEBOEK	2667.21	106	ED	1096.87	181	GILLES	653.54	256	SEARCH	487.81
32	LEXICOGRAPHICAL	2641.68	107	COPULATIVE	1091.27	182	AS	652.44	257	SENSES	486.42
33	EQUIVALENTS	2488.19	108	SUBLEXICAL	1089.83	183	WÖRTERBUCH	646.58	258	LA	485.61
34	USERS	2484.20	109	HEADWORD	1087.33	184	ENTRY	644.56	259	COMPONENT	483.39
35	SOUTH	2450.55	110	SPECIFIC	1087.08	185	BERGENHOLTZ	636.47	260	OUTER	483.31
36	INFORMATION	2409.82	111	ACCESS	1078.63	186	EDITORS	634.70	261	GRAMMAR	481.12
37	LEXICOGRAPHERS	2374.74	112	ITEM	1062.15	187	ZGUSTA	626.37	262	SAGER	474.41
38	SEMANTIC	2366.90	113	SA	1060.79	188	KEYWORDS	623.40	263	VARIOUS	473.96
39	TEXTS	2346.14	114	VERBS	1054.38	189	ISIXHOSA	616.27	264	SEMANTICALLY	473.43
40	CORPORA	2335.47	115	SLANG	1042.87	190	REGARD	609.50	265	GE	472.78
41	VERB	2335.04	116	EXAMPLE	1011.57	191	SINGULAR	609.03	266	DESCRIPTIVE	470.10
42	LEXICOGRAPHER	2270.21	117	STELLENBOSCH	1005.08	192	ISIZULU	606.17	267	TERMINOLOGIE	465.12
43	NOUN	2218.57	118	MACROSTRUCTURE	1000.16	193	LEXEMES	603.77	268	ALLEX	464.73
44	LEXIKOS	2212.56	119	FRAÇAIS	990.08	194	AFRIKA	602.14	269	MEDIOSTRUCTURE	464.73
45	TERMS	2187.85	120	COPYRIGHT	972.04	195	NIEMEYER	602.10	270	METALEXICOGRAPHICAL	464.73
46	MONOLINGUAL	2158.57	121	COMPILERS	959.51	196	KRIEL	599.05	271	COLLOCATION	462.68
47	MEANING	2142.08	122	STRUCTURE	956.05	197	FREQUENCY	595.98	272	PRESENTED	459.78
48	KA	2138.15	123	LEARNER'S	955.96	198	HARTMANN	592.92	273	GROOT	458.70
49	GOUWS	2121.63	124	CONSONANT	949.92	199	TYPES	591.36	274	TECHNICAL	458.32
50	PRETORIA	2114.45	125	TERMINOLOGISTS	949.67	200	IDIOMS	589.38	275	NORTHERN	456.66
51	LEMMATA	2107.06	126	ENCYCLOPEDIA	933.89	201	CROSS	588.67	276	IN	455.25
52	EDS	2058.02	127	METALEXICOGRAPHY	929.46	202	MANTŠU	585.96	277	TRANSLATIONS	454.35
53	TARGET	2057.42	128	USE	929.07	203	TARP	585.90	278	PRINCIPLES	454.32
54	LEXICON	2010.31	129	LEXIKOGRAFIE	919.36	204	NON	585.67	279	VENDA	454.28
55	LEMMAS	2004.32	130	PROJECT	915.25	205	SYNONYMY	583.33	280	SENTENCES	452.24
56	AFRICA	1954.78	131	DSAE	909.25	206	SEMANTICS	581.43	281	ENCYCLOPAEDIC	451.08
57	EXAMPLES	1954.06	132	YA	897.54	207	CULTURAL	578.34	282	ETYMOLOGY	451.08
58	ARTICLES	1828.75	133	INDIGENOUS	889.38	208	ABBREVIATIONS	576.14	283	FEEDBACK	450.86
59	REFERENCE	1806.51	134	ARTIKEL	889.05	209	AFRIKATALE	575.86	284	THOUGHTFUL	449.75
60	USAGE	1777.37	135	OF	880.04	210	CLUVER	575.86	285	NICHE	445.59
61	VOCABULARY	1774.09	136	ARABIC	879.79	211	DURAMAZWI	575.86	286	KONERDING	444.52
62	COMPILATION	1714.60	137	LSP	869.89	212	ORDERING	574.86	287	ISN	444.37
63	SPEAKERS	1688.41	138	COBUILD	850.31	213	MULTI	567.96	288	ENCYCLOPEDIA	442.46
64	NICHING	1677.09	139	DATA	836.42	214	ALSO	562.00	289	SWEDISH	442.12
65	PRINSLOO	1665.49	140	BANTU	834.20	215	RESEARCH	559.41	290	ACCORDING	441.38
66	ZULU	1648.46	141	LEKSIKOGRAFIE	808.22	216	LOUW	557.01	291	BUSINESS	440.43
67	MULTILINGUAL	1641.82	142	MACROSTRUCTURAL	808.22	217	DIALECTS	555.02	292	PLURALS	440.21
68	COLLOCATIONS	1631.31	143	EQUIVALENT	808.09	218	PREFIX	549.19	293	THEREFORE	439.82
69	DEFINITIONS	1606.27	144	COPULATIVES	798.12	219	SL	545.88	294	NEEDS	439.36
70	KE	1597.31	145	SYNONYM	796.57	220	INL	545.55	295	REFERENCING	438.69
71	AFRILEX	1596.26	146	USED	795.13	221	LOUWRENS	545.55	296	LEMMATISATION	436.90
72	MORUTSI	1586.16	148	NASAL	792.15	222	WYK	541.63	297	TREATMENT	436.57
73	ENTRIES	1579.21	147	IS	791.62	223	PRONUNCIATION	538.62	298	KAF	434.94
74	LINGUISTICS	1558.84	149	ETC	789.05	224	DGC	538.35	299	SEGMENT	434.49
75	HAUSMANN	1543.78	150	CILUBÀ	788.02	225	ROM	538.23	300	VERKLARENDE	434.42

Addendum 3: Keywords and keyness values for *Lexikos* sub-corpus 3: 2003–2008

#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness	#	Keyword	Keyness
1	DICTIONARY	60677.92	76	SETSWANA	1778.25	151	SCHOLTZ	983.35	226	HARARE	697.04
2	DICTIONARIES	34268.95	77	FRANÇAIS	1771.51	152	ISIXHOSA	982.02	227	MEDIOSTRUCTURE	693.19
3	LANGUAGE	22342.33	78	KA	1759.05	153	ADJ	981.19	228	SPOKEN	691.46
4	LEXICOGRAPHY	16604.31	79	USAGE	1737.13	154	SUFFIX	980.06	229	DEUTSCH	690.33
5	LANGUAGES	11389.05	80	WOORDEBOEKE	1723.37	155	DEFINITIONAL	973.51	231	HAUSA	685.63
6	ENGLISH	9489.79	81	ITEM	1678.11	156	LEKSIKOGRAFIE	972.39	230	EVALUATION	684.54
7	LEXICOGRAPHIC	8408.29	82	INTERNET	1649.28	157	COMPILERS	971.13	232	BOSMAN	683.80
8	SHONA	8319.34	83	STELLENBOSCH	1641.89	158	GRAMMAR	968.81	233	SVEN	683.58
9	CORPUS	8222.76	84	GABON	1631.54	159	FREQUENCY	962.22	234	EENTALIGE	683.56
10	NDEBELE	7895.39	85	ORTHOGRAPHY	1628.97	160	LEKSIKOGRAFIE	953.14	235	MACROSTRUCTURAL	683.56
11	USER	7256.95	86	EQUIVALENCE	1626.88	161	SHANGANI	953.14	236	CROSS	680.86
12	BILINGUAL	6939.91	87	ISN	1624.42	162	SPELLING	949.29	237	CAPE	680.60
13	LEXICAL	6218.81	88	TÜBINGEN	1617.46	163	YA	944.02	238	FIELD	674.58
14	AFRIKAANS	6101.93	89	NON	1614.95	164	TSHIVENDA	943.51	239	TERMINOLOGISTS	673.93
15	WORD	6093.01	90	AFRICA	1570.08	166	CULTURE	934.00	240	KISWAHILI	669.92
16	WORDS	5703.51	91	SA	1569.55	165	DIALECT	926.54	241	PREFIX	665.35
17	AFRICAN	5608.01	92	ET	1569.43	168	GRUYTER	913.13	242	COPULATIVE	655.69
18	USERS	5540.67	93	LIST	1545.90	169	ADJECTIVES	912.92	243	NIENABER	655.69
19	MONOLINGUAL	5325.01	94	AFRIKAANSE	1530.81	167	INDIGENOUS	906.86	244	HÜLLEN	654.68
20	LEXICOGRAPHERS	5118.73	95	NIEMEYER	1529.75	170	LEARNER	905.73	245	VERBAL	651.68
21	LEMMA	5052.34	96	HARTMANN	1518.58	171	SEPEDI	905.00	246	TREATMENT	649.34
22	TEXT	4614.96	97	TSHWANELEX	1482.67	172	MICROSTRUCTURAL	904.76	247	NA	642.26
23	SOTHO	4607.16	98	TERMINOLOGICAL	1481.61	173	LEMMATISATION	900.28	248	COMMUNICATION	638.75
24	TRANSLATION	4161.60	99	NOUN	1481.39	174	ETYMOLOGICAL	898.83	249	IS	638.30
25	ARTICLE	4062.62	100	TAAI	1462.62	175	HTTP	895.37	250	PRESCRIPTIVE	634.44
26	LEXICOGRAPHICAL	3968.11	101	OUTER	1459.17	176	YILUMBU	895.37	251	RESEARCH	628.99
27	LEXICOGRAPHER	3895.24	102	COLLOCATIONS	1441.60	177	PRONUNCIATION	892.31	252	POLYSEMOUS	628.43
28	LEARNERS	3806.67	103	PHAROS	1415.66	178	LEKSIKOGRAFIESE	885.74	253	MAIOR	627.43
29	LINGUISTIC	3735.64	104	VENDA	1409.96	179	EXPRESSIONS	884.43	254	LA	627.17
30	TEXTS	3717.04	105	DURAMAZWI	1396.02	180	ENTRY	883.62	255	ISICHAZAMAZWI	625.79
31	EQUIVALENTS	3552.15	106	HEADWORD	1383.63	181	ADJECTIVE	870.45	256	INTERLINGUAL	625.04
32	REFERENCE	3425.34	107	NIE	1380.57	182	RECHISHONA	866.49	257	SUBLEXICAL	625.04
33	GOUWS	3321.64	108	IDIOMS	1332.29	183	LSP	866.42	258	SENSES	622.08
34	ONLINE	3174.96	109	GRAMMATICAL	1328.31	184	INCLUSION	862.09	259	ELECTRONIC	621.15
35	SCHRYVER	3163.63	110	TERM	1327.90	185	THESE	859.80	260	ZIMBABWEAN	618.98
36	WIEGAND	3051.72	111	MEANINGS	1321.60	186	SEMANTICS	855.74	261	INFLECTIONAL	618.57
37	SWAHILI	3051.10	112	EXAMPLE	1314.65	187	ZGUSTA	847.23	262	X	616.54
38	SESOTHO	3029.74	113	ADVERBS	1314.56	188	TYPES	845.42	263	YIPUNU	616.17
39	WOORDEBOEK	2994.28	114	USE	1297.21	189	MICROSTRUCTURE	844.61	264	CRITERIA	611.04
40	LEXIKOS	2965.39	115	TARGET	1286.53	190	OF	839.19	265	TRILINGUAL	608.26
41	MEANING	2920.72	116	SPECIFIC	1270.07	191	COMPILING	835.00	266	MAROLE	606.54
42	LEARNER'S	2912.65	117	DEFINITION	1250.81	192	ALPHABETICAL	834.15	267	WÖRTERBÜCHER	606.54
43	WAT	2895.51	118	SOUTH	1239.26	193	CL	830.74	268	EDITORS	606.11
44	EDS	2797.59	119	SYNONYMS	1234.83	194	LONGMAN	829.72	269	NGUNI	605.85
45	ITEMS	2712.20	120	ARTIKEL	1232.34	195	PRESENTED	826.68	270	ETYMOLOGY	603.59
46	TERMINOLOGY	2686.55	121	HESSING	1232.34	196	PHONETIC	824.18	271	BASED	601.95
47	COMPILATION	2645.66	122	HEADWORDS	1221.03	197	PRON	820.91	272	PRESENTATION	601.39
48	VOCABULARY	2605.56	123	NOUNS	1219.52	198	EURALEX	818.35	273	TOIT	599.28
49	DATA	2571.41	124	REFERENCES	1217.97	199	CULTURAL	809.36	274	AS	598.02
50	DEFINITIONS	2558.94	125	PHILIPPSON	1213.09	200	MAVOUNGOU	808.72	275	GA	596.57
51	LEBOA	2512.87	126	FORMS	1192.56	201	KEYWORDS	808.70	276	FOREIGN	593.24
52	SEMANTIC	2478.52	127	VERBS	1191.18	202	REGARD	793.71	277	DANISH	591.96
53	GABONESE	2472.23	128	MACROSTRUCTURE	1189.81	203	VARIANTS	791.83	278	ACCESS	587.32
54	BERGENHOLTZ	2464.73	129	LEXICON	1188.27	204	EDITION	787.19	279	SHOULD	585.16
55	ZULU	2401.48	130	EQUIVALENT	1180.09	205	KHWE	779.84	280	LINGUISTS	585.06
56	INFORMATION	2372.10	131	USED	1175.08	206	LEMMATIZATION	760.58	281	ZEZURU	577.66
57	ENTRIES	2360.37	132	AL	1175.01	207	TWEETALIGE	760.58	282	ALLEX	568.03
58	ED	2282.52	133	BANTU	1169.46	208	LEXSM	750.95	283	LEMMATIZED	568.03
59	CF	2236.65	134	DICTIONNAIRE	1164.80	209	POS	748.93	284	BIBLIOGRAPHY	565.35
60	ARTICLES	2217.55	135	OXFORD	1120.67	210	STRUCTURE	748.79	285	THEORETICAL	561.91
61	LINGUISTICS	2198.63	136	TLS	1113.15	211	DIFFERENT	746.43	286	FORM	555.74
62	SPEAKERS	2153.77	137	GILLES	1105.95	212	CANADIAN	743.56	287	ORTHOGRAPHIC	553.87
63	PRETORIA	2145.64	138	LE	1096.98	213	KNOWLEDGE	741.92	288	METALANGUAGE	551.73
64	TARP	2134.20	139	FUNCTIONS	1090.11	214	WA	741.80	289	LEXICOGRAPHIE	548.77
65	PRINSLOO	2133.77	140	LEXICOGRAPHICA	1087.93	215	ETC	734.21	290	RAPONDA	548.77
66	EXAMPLES	2094.41	141	COBUILD	1083.54	216	DESCRIPTION	730.45	291	MAURICE	547.90
67	CORPORA	2061.21	142	MORPHOLOGICAL	1072.77	217	MULTILINGUAL	719.94	292	HENNING	545.84
68	UNIVERSITY	2043.17	143	DEM	1068.77	218	SPEECH	717.52	293	JOFFE	544.51
69	LEMMAS	2010.79	144	LEMMATA	1059.53	219	PROVERBS	715.76	294	NEEDS	539.75
70	DIALECTS	1970.74	145	COMPILED	1057.86	220	HADIBE	711.44	295	INTRALINGUAL	539.15
71	AFRILEX	1944.81	146	WÖRTERBUCH	1049.42	221	ADDENDUM	711.21	296	METALEXICOGRAPHY	539.15
72	TERMS	1884.03	148	HAUSMANN	1047.65	222	SEARCH	709.00	297	MULTILEXICAL	539.15
73	ZIMBABWE	1841.84	147	KE	1029.89	223	ALSO	708.81	298	CATEGORIES	535.78
74	ALRI	1819.67	149	EFL	1011.61	224	LABELS	705.25	299	JOURNAL	532.78
75	VERB	1791.75	150	CHIMHUNDU	991.65	225	FUNCTION	700.96	300	SUCH	532.59